Education in Bhutan
Findings from Bhutan’s experience in PISA for Development
Education in Bhutan

Findings from Bhutan’s experience in PISA for Development

National Project Centre
Bhutan Council for School Examinations and Assessment
Bhutan’s PISA-D National Report represents a multi-team effort of various national and international organizations. The National Project Centre is grateful to the Ministry of Education for initiating Bhutan’s participation in PISA-D 2017 and for providing financial support for the project implementation.

We would like to express our gratitude to the members of PISA-D National Steering Committee, Central Level Core Group, Dzongkhag Level Core Group and School Level Core Group for planning and coordinating the implementation of the project. The project would not have been successful without the cooperation and support from students, teachers and principals in the field.

We would like to acknowledge officials from Ministry of Education, Royal Education Council and Bhutan Council for School Examinations and Assessment for the support rendered in the test administration process.

We would also like to extend our sincere thanks to the Ministry of Education, Royal Education Council, Paro College of Education and Bhutan Council for School Examinations and Assessment for their feedbacks on the PISA-D National Report.

Apart from being thankful to the OECD, ETS and Westat for the technical support and guidance provided at various stages of the project, we are also grateful for designing the template for Bhutan PISA-D National Report and for finalizing it.
Findings indicate students in Bhutan in general ...

- have higher success rates in items requiring lower cognitive skills, however, there is a significant gap in performance in more demanding tasks.
- demonstrate relatively broader knowledge and understanding compared to other PISA-D countries.
- have performed at par with top PISA-D countries but significantly below OECD average.
# CONTENTS

## CHAPTER 1  
**PISA AND PISA FOR DEVELOPMENT**

- WHAT IS PISA? 2
- THE PISA ASSESSMENT 2
- PISA FOR DEVELOPMENT 3
- THE PISA-D TEST 3
- CAPACITY BUILDING 4
- PARTICIPATING COUNTRIES 4
- WHY BHUTAN IS PARTICIPATING IN PISA-D 4

## CHAPTER 2  
**REPORTING OF RESULTS**

- ANALYSES OF TEST LENGTH AND OF ITEM FUNCTIONING IN BHUTAN 10
- OVERALL PERFORMANCE OF BHUTAN IN PISA-D 10
- READING LITERACY ASSESSMENT IN PISA-D 12
  - SOLUTION RATES IN READING IN BHUTAN 13
  - OVERALL PERFORMANCE AND GENDER DIFFERENCES 13
  - RELATIVE STRENGTHS 15
  - RELATIVE WEAKNESSES 17
  - ANALYSIS OF ILLUSTRATING ITEMS 18
- MATHEMATICAL LITERACY ASSESSMENT IN PISA-D 23
  - SOLUTION RATES IN MATHEMATICS IN BHUTAN 24
  - OVERALL PERFORMANCE AND GENDER DIFFERENCES 24
  - RELATIVE STRENGTHS 26
  - RELATIVE WEAKNESSES 27
  - ANALYSIS OF ILLUSTRATING ITEMS 29
- SCIENTIFIC LITERACY ASSESSMENT IN PISA-D 33
  - SOLUTION RATES IN SCIENCE IN BHUTAN 34
  - OVERALL PERFORMANCE AND GENDER DIFFERENCES 34
  - RELATIVE STRENGTHS 36
  - RELATIVE WEAKNESSES 37
  - ANALYSIS OF ILLUSTRATING ITEMS 38

## CHAPTER 3  
**SUMMARY AND SUGGESTIONS FOR BHUTAN’S FORTHCOMING PARTICIPATION IN PISA 2021**

- SUMMARY OF FINDINGS 45
  - OVERALL PERFORMANCE 45
  - SUMMARY OF RELATIVE STRENGTHS IN THE THREE DOMAINS 45
  - SUMMARY OF RELATIVE WEAKNESSES IN THE THREE DOMAINS 46
  - IMPLICATIONS OF THE STRENGTHS AND WEAKNESSES FOR BHUTAN 47
  - NEXT STEPS IN BHUTAN’S PARTICIPATION IN PISA 2021 49
The Ministry of Education’s purpose is to shape an education system for our country that delivers equitable and excellent outcomes for all of our children and young people. A strong focus on student learning and well-being underpins all our policy and the services we provide.

It is to help us achieve our Ministry’s purpose that we joined the OECD’s Programme for International Student Assessment for Development, PISA-D. This programme aims to evaluate education systems in the low and middle-income countries by assessing the extent to which 15-year-old students, near the end of their compulsory education, have acquired key knowledge and skills that are essential for full participation in modern societies. Bhutan’s participation in PISA-D demonstrates the importance we place on the educational achievement of our children and young people.

In this report, a team of Ministry of Education and Bhutan Council for School Examinations and Assessment officials have collated and analysed the information from our participation in PISA-D so that it can be used to benefit the education sector and, therefore, the children in our education system. This report contributes sound data, information and analysis for work undertaken to support the Government’s existing education policies and our education policies, strategies and programmes in the future. This opportunity for international comparison and international learning is an extremely valuable aspect of our participation in PISA-D.

The PISA-D assessment focuses on the core school subjects of reading, mathematics and science, and does not just ascertain whether students can reproduce knowledge; it also examines how well students can extrapolate from what they have learned and can apply that knowledge in unfamiliar settings, both in and outside of school.

You will find in the second chapter of this report a detailed and thorough analysis of what PISA-D data tells us about our students’ performances in reading, mathematics and science, how this performance compares to students in other countries. In the final chapter, the report lays out the policy implications of the findings and results and points the way for strengthening those of our current education policies that are most relevant to us.

The most important thing that PISA-D data tell us about our education system is that Bhutanese 15-year-olds performed at par with 15-year-olds of top PISA-D countries but lags significantly behind that of OECD countries and of the best education systems in Asia. This is shown by the average solution rates: 45.3% in reading literacy, 38.8% in mathematics literacy and 45.1% in scientific literacy. On an average in Bhutan, boys outperformed girls in mathematics and girls outperformed boys in reading. There was no significant difference between boys and girls in science solution rates. These are important messages for us regarding the quality of student learning and equity of our education system.

Ministry of Education, Royal Government of Bhutan intend to respond fully to the findings and messages contained in this report and to follow up the suggestions regarding effective interventions that are set out in the final chapter. The effective interventions highlighted in this report include actions designed to improve educational outcomes and the quality of instruction.
Success in education relies on many people and organisations across the community working together for the benefit of children and young people. We trust the information in this report will help all of us involved in improving our education system.

I would also like to commend our Bhutan PISA-D team, especially the National Project Manager, Tenzin Dorji (Ex. Secretary, BCSEA), Deputy National Project Manager and Mathematical Literacy Lead Coder, Arjun Kumar Gurung, Reading Literacy Lead Coder, Kinley Dema, Scientific Literacy Lead Coder, Sonam Lhamo, PISA-D Focal Officer, Pem Tshering and our Lead Analyst, Karma Jigme Lepcha, the principal author of this report.

(Jai Bir Rai)
Minister
Ministry of Education
Bhutan
The purpose of this publication, the first of its kind in Bhutan, is to present the results of the country’s participation in the OECD’s Programme for International Student Assessment for Development (PISA-D). PISA has become the world’s premier yardstick for evaluating the quality, equity and efficiency of school systems and it is a mark of our progress in education that Bhutan is participating in it. This report describes the results achieved by our students, the outcome of resources invested in our education system, and the learning environments in our schools and communities, in ways that allow for comparisons with other countries participating in PISA-D, some of them from our own region. The data and analysis contained in these pages will help our government and our educators identify the main challenges for education policy in Bhutan, and can inform the development of effective strategies and policies to confront them.

Bhutan joins almost 90 countries and economies that have participated in PISA since it began in 2000. Bhutan is one of the nine countries that partnered with the OECD through the “PISA for Development” initiative, whose aim is to make PISA more accessible and relevant to middle- and low-income countries like ours. In respect of Bhutan, an important enhancement in this initiative concerned the PISA cognitive assessment instruments themselves, which were re-designed to capture a wider range of performance levels, but on the same scales as those used in the regular PISA assessment. We have also benefited from the capacity development that has been built into the PISA for Development project and this will be utilised by us in future cycles of PISA as well as in our own national assessments.

Bhutan’s participation in PISA for Development would not have been possible without the contributions of Ministry of Education, Bhutan Council for School Examinations and Assessment and Royal Education Council.
<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>BCSEA</td>
<td>Bhutan Council for School Examinations and Assessment</td>
</tr>
<tr>
<td>BTN</td>
<td>Bhutan</td>
</tr>
<tr>
<td>DEO</td>
<td>District Education Officer</td>
</tr>
<tr>
<td>DSE</td>
<td>Department of School Education</td>
</tr>
<tr>
<td>DOM</td>
<td>Dominican Republic</td>
</tr>
<tr>
<td>ECU</td>
<td>Ecuador</td>
</tr>
<tr>
<td>EMSSD</td>
<td>Education Monitoring and Support Service Division</td>
</tr>
<tr>
<td>GTM</td>
<td>Guatemala</td>
</tr>
<tr>
<td>HND</td>
<td>Honduras</td>
</tr>
<tr>
<td>HRC</td>
<td>Human Resource Committee</td>
</tr>
<tr>
<td>HRD</td>
<td>Human Resource Division</td>
</tr>
<tr>
<td>ICT</td>
<td>Information and Communication Technology</td>
</tr>
<tr>
<td>KHM</td>
<td>Cambodia</td>
</tr>
<tr>
<td>MoE</td>
<td>Ministry of Education</td>
</tr>
<tr>
<td>NEA</td>
<td>National Education Assessment</td>
</tr>
<tr>
<td>NPC</td>
<td>National Project Centre</td>
</tr>
<tr>
<td>OECD</td>
<td>Organisation for Economic Co-operation and Development</td>
</tr>
<tr>
<td>PISA</td>
<td>Programme for International Student Assessment</td>
</tr>
<tr>
<td>PISA-D</td>
<td>PISA for Development</td>
</tr>
<tr>
<td>PPD</td>
<td>Policy and Planning Division</td>
</tr>
<tr>
<td>PRY</td>
<td>Paraguay</td>
</tr>
<tr>
<td>REC</td>
<td>Royal Education Council</td>
</tr>
<tr>
<td>RUB</td>
<td>Royal University of Bhutan</td>
</tr>
<tr>
<td>SDG</td>
<td>Education Sustainable Development Goal</td>
</tr>
<tr>
<td>SEN</td>
<td>Senegal</td>
</tr>
<tr>
<td>SGP</td>
<td>Singapore</td>
</tr>
<tr>
<td>SWE</td>
<td>Sweden</td>
</tr>
<tr>
<td>TEO</td>
<td>Thromdey Education Officer</td>
</tr>
<tr>
<td>ZMB</td>
<td>Zambia</td>
</tr>
</tbody>
</table>
Bhutan in PISA-D
CHAPTER 1  PISA AND PISA FOR DEVELOPMENT

In Bhutan during the month of November 2017 more than 2,400 15-year-old students in 7th grade or above from 53 randomly selected schools across the country took a two-hour test in reading, mathematics and science. These tests were not directly linked to Bhutan’s school curriculum – rather, they were competency based and internationally comparable. These tests were designed by the Organisation for Economic Co-operation and Development (OECD) to assess the extent to which students at the end of compulsory education can apply their knowledge to real-life situations and be equipped for full participation in society. These tests are part of an international large-scale assessment of learning which is managed by the OECD and is called the Programme for International Student Assessment or PISA for short.

This first ever participation of Bhutan in PISA was intended as a pilot testing, to prepare the country to take part in future rounds of PISA, starting with PISA 2021. The main goals were to confirm that the test could provide reliable and valid results about Bhutan’s students, and that the procedures ensured full comparability of the results with those of other countries. The main conclusions that can be drawn from this first participation, therefore, informed: whether the Bhutanese institutions have the capacity to carry out an international assessment; how appropriate the PISA test is as a measurement of learning outcomes in Bhutan; and what main difficulties are encountered by Bhutanese students when answering the PISA test. However, due to small sample of schools and students that took part in the PISA pilot in Bhutan relative to the normal PISA and PISA-D sample, and because Bhutan was not required to implement the full technical standards for this first pilot participation, the results have not been presented on the internationally comparable PISA scale.

PISA evaluates students studying in class VII (7th grade) and above who are aged between fifteen years and three months and sixteen years and two months at the time of the evaluation. The assessment focuses on the core domains of reading, mathematics and science. The assessment does not just ascertain whether students can reproduce knowledge rather it examines how well students can extrapolate from what they have learned and can apply that knowledge in unfamiliar settings, both in and outside of school. This approach reflects the fact that modern economies reward individuals not for what they know, but for what they can do with what they know.

In Bhutan the PISA pilot test was administrated between 1st and 15th November 2017. The sample of schools was selected by the OECD, the international organization in charge of the study, based on a complete list of all schools with eligible students in the country submitted by national authorities, and of complete listings of 15-year-old students in these schools submitted by the school administrators in the selected schools. This data is not representative of the entire population of 15-year-old students in the country but it does provide a reliable assessment of the strengths and weaknesses among students in Bhutan in the three domains tested.

As of 2017, Bhutan had a total of 515 schools, consisting of 310 primary schools, 72 lower secondary schools, 72 middle secondary schools and 61 higher secondary schools. The total enrolment from Pre-primary through to class XII was 168,092 (Annual Education
Statistics 2017) and 10,725 of these students were estimated to be 15 years old (EMIS 2017).

The Bhutan sample consisted 2,457 students belonging to 53 schools. This sample included establishments of all the dependencies and modalities existing in the country, as well as of all the regions, in urban and rural areas. In each establishment, 15 to 70 students of 15 years were randomly selected.

**What is PISA?**

Launched by the OECD in 1997, PISA assesses 15-year-olds’ proficiency in reading, mathematics and science and measures students’ skills in applying what they have learned in school to real-life situations. PISA cycles have been completed in 2000, 2003, 2006, 2009, 2012 and 2015 and the 2018 cycle is under way. PISA is an on-going programme that offers insights for education policy and practice, and that helps to monitor trends in students’ acquisition of knowledge and skills across countries and in different demographic subgroups within each country. Through PISA results, policy makers can gauge the knowledge and skills of students in their own countries in comparison with those in other countries, set policy targets against measurable goals achieved in other education systems, and learn from policies and practices of countries which have demonstrated improvement. This kind of international benchmarking is more relevant now than ever, given that every country in the world has signed up to the Education Sustainable Development Goal (SDG) agenda which is about ensuring that every child and young person achieves at least basic levels of proficiency in reading and mathematics.

**The PISA assessment**

PISA facilitates international comparison of countries’ education systems through the use of common items, used by all participating countries, which are all located on a common measurement scale. However, before the results of a new participant in PISA are reported on the international PISA scale, PISA verifies that these common items do maintain their expected measurement properties when administered to students in the country. If this condition is met, student responses can be transformed, using advanced statistical techniques, into PISA scores that are located along specific scales developed for each subject area. These scales are divided into levels that represent groups of PISA test questions, beginning at Level 1 with questions that require only the most basic skills to complete and increasing in difficulty with each level up to 6. For example, a student who lacks the skills needed to complete the easiest questions on a PISA test correctly would be classified as below Level 1, while a student who has these skills would be at a higher level.

If the sample of students is representative of a country's internationally agreed target population for PISA, a country's mean score can be computed as the average of all student scores in that country. PISA mean scores can be used to rank participating countries according to their performance in reading, mathematics and science. PISA does not give a collective score for all subjects combined; rather it gives a score for each subject area and this can be used to determine rankings by the mean score of each area. Because the
sample for Bhutan was not designed to provide representative results for the country, the results of the first pilot participation of Bhutan refer, strictly speaking, only to the sample that took part in PISA-D, and caution is required when conclusions are drawn from these results about the whole country.

PISA is an ongoing programme that, over the longer term, will lead to the development of a body of information for monitoring trends in the knowledge and skills of students in various countries as well as in different demographic subgroups of each country. Policy makers around the world use PISA findings to gauge the knowledge and skills of students in their own country/economy in comparison with those in other participating countries/economies, establish benchmarks for improvements in the education provided and/or in learning outcomes, and understand the relative strengths and weaknesses of their own education systems.

PISA for Development

Over the past two decades, PISA has steadily increased the number of participating countries, from 44 in 2000 to 82 in 2018. As the number of countries joining PISA increases, PISA evolves to successfully cater for a larger and more diverse group of participants. Bhutan decided to join the PISA for Development (PISA-D) project in the beginning of 2017. The project is also a contribution to the monitoring of international educational targets related to the Education Sustainable Development Goal (SDG), adopted by the United Nations General Assembly in 2015 as part of the Agenda for Sustainable Development. To accomplish its aims, the PISA-D project sets out to increase the resolution of the PISA tests at the lower end of the student performance distribution.

The lessons learned from PISA-D are being incorporated into the regular PISA cycles, starting with the 2021 cycle and Bhutan is already benefiting from these lessons as it prepares for its participation in the main PISA assessment in 2021.

The PISA-D test

The PISA-D school-based assessment is a two-hour test that students complete with pencil and paper. The test includes a combination of questions from the domains of reading, mathematics and science. Each student was given one of 12 possible test booklets, which overlap in content. By administering different booklets to different students, PISA-D can measure a wide range of knowledge and skills at the country level, without the need to administer an exceedingly long and complex test to individual students. All test booklets administered in PISA-D contain items that were part of the PISA 2015 instruments, to ensure that results can be compared with those of countries that participated in PISA 2015.

Each test booklet is completed by a sufficient number of students to make appropriate estimates of the solution rates on all items by students in each country and in relevant subgroups within a country (such as boys and girls). Just as PISA, however, PISA-D is not designed to estimate the performance of individual students or schools: its results are most valid and reliable when aggregated across a sufficient number of students. Comparability with PISA 2015, which was administered both on paper and on computers, is assured through common items.
While PISA-D has been implemented within the overall PISA framework and in accordance with PISA’s technical standards and usual practices, it includes new features and enhancements to make the assessment more accessible and relevant to middle- and low-income countries. With regard to the test, these features and enhancements include:

- an equal treatment of the three major domains tested - reading, mathematics and science - unlike PISA, where one of the domains is given a particular focus in each cycle;
- test instruments that cover a wider range of performance at the lower levels of proficiency, while still providing scores that cover the whole of the PISA framework and are comparable to the main PISA results; and
- modified test instruments that have a reduced reading burden, in recognition of the lower levels of reading literacy capacity in middle- and low-income countries.

Capacity building

A further feature unique to PISA-D is the learning and capacity-building opportunities that have been built into each phase of project implementation. Bhutan has benefited from these capacity building opportunities and will continue to receive capacity building inputs through its participation in PISA 2021.

Participating countries

The PISA-D project has been carried out by the OECD in partnership with Bhutan and eight other countries: Cambodia, Ecuador, Guatemala, Honduras, Paraguay, Panama, Senegal and Zambia.

Why Bhutan is participating in PISA-D

The main reason why Bhutan is participated in PISA-D is to gain experience from this initiative to help it prepare for participation in the PISA 2021 cycle. One of the main reasons Bhutan wants to participate in PISA 2021 is because of its policy makers’ wish to understand how the performance of students in the country compares, in relation to international benchmarks and to countries facing similar challenges elsewhere, and to identify the factors that are associated with under performance in order to effectively eliminate it. The PISA-D results contained in this report provide these policy makers with some data and evidence that can be used to determine what they can do to improve Bhutan’s education system. However, it is through its participation in PISA 2021 that the country will gain the fullest insights that will help it to ensure ultimately that its students obtain the skills needed to succeed in tomorrow’s world and as set out in the Education Sustainable Development Goal Framework adopted by the United Nations in 2015.

All countries are committed to achieving the key Education SDG target of all children and young people reaching at least minimum levels of proficiency in reading and mathematics by 2030. In Bhutan, this means ensuring all young citizens have the knowledge, skills and capabilities necessary to achieve their full potential, contribute to an increasingly
interconnected world, and live a fulfilling life. Every student in Bhutan should have the opportunity to achieve excellence in learning outcomes that not only reflect the wisdom of the country's dynamic culture and traditions but are also comparable to student outcomes in high-performing international education systems.
Achievement and attainment outcomes at 15 in Bhutan
CHAPTER 2 REPORTING OF RESULTS

This national report and other related communication products present Bhutan’s results in context to the countries that participated in PISA 2015 and PISA-D and include relevant analyses and information based on the policy priorities of Bhutan. This report constitutes a summary of key results and analysis designed to stimulate a constructive debate on improvement, building upon and enriching already existing data and evidence from national, regional or international sources. This national report should be considered as a starting point for a discussion of the results and implications for policy with a range of stakeholders, including students, parents, teachers, school principals, academia, civil society, media and central and local government.

The report presents analysis of the performance of the students in Bhutan in reading, mathematics and science items that were administered in the country as part of the PISA-D assessment. Bhutan’s performance will be compared both with the PISA 2015 and PISA-D averages and with selected OECD and PISA-D countries. Five countries have been chosen for comparison with Bhutan: the two highest-performing PISA-D countries and three countries (one high, one middle and one low performing country) which participated in PISA 2015. These five countries will be called the “reference countries”. The specific relative strengths and weaknesses of Bhutan’s 15-year-olds will be identified; that is, items where the students from Bhutan performed unexpectedly well or unexpectedly badly compared to their overall distance from the PISA-D average or from the reference countries. Altogether, there are 38 such so-called “conspicuous items”: 10 in reading (5 for strengths and 5 for weaknesses); 15 in mathematics (7 for strengths and 8 for weaknesses); and 13 in science (6 for strengths and 7 for weaknesses). The report focuses in particular on these conspicuous items.

Different approaches have been followed in compiling this report in order to identify relative strengths and relative weaknesses of Bhutan compared to the average of all the PISA-D countries and compared to the reference countries. To confirm that the items of the PISA test maintained their expected measurement properties when administered to students in Bhutan, and no bias would result from errors in adaptation or due to cultural sensitivities, various differential-item-functioning (DIF) statistics were analysed (this statistical review follows a more qualitative, and substantive review by every country participating in PISA).

Once the appropriateness of the PISA test was statistically confirmed, the average solution rates across items were examined to describe the performance of the students who took the PISA test in Bhutan in comparison to the PISA-D average and to the average solution rates of reference countries. Finally, the naïve “percent-correct” measure for solution rates was transformed on a scale – the so-called delta scale – that is more adequate for comparisons across items of varying difficulty. These “delta scores” (described later in this Chapter) were then compared across items and countries to identify a small set of “conspicuous items” and to draw preliminary conclusions about strengths and weaknesses of students in Bhutan which can inform improvement efforts, including teacher training and the design of national curricula and assessments.

The report concludes with some suggestions for further analysis and actions that may be pursued through Bhutan’s forthcoming participation in PISA 2021.
Analyses of test length and of item functioning in Bhutan

The report presents analysis of test length and item functioning in Bhutan by looking into the omission rates, not reached items and item fit statistics. The patterns observed among Bhutan’s students are typical of those of middle-income countries participating in PISA.

On a number of relatively difficult items requiring open-ended responses (such as numeric results in mathematics or words/sentences in reading and science), the omission rates are above 10% - meaning that more than one out of ten students did not attempt to answer these items.

A small but significant fraction of students did not finish the test, perhaps because of insufficient fluency in reading; this pattern is indeed most notable among students who were assigned (at random) to a reading test.

While the vast majority of items included in the PISA test did show similar measurement properties as in other countries – retaining, in particular, the ability to discriminate between more and less proficient students, with similar strength as in other countries on average – a few items exhibit small deviations from international patterns in their difficulty threshold, meaning that they appear to be somewhat easier or somewhat more difficult than expected for Bhutan’s students. The small number of such items (less than 20% in each domain), and the general conclusion that even these items nevertheless support the construction of a coherent proficiency scale, leads us to conclude that the PISA test in Bhutan is appropriate to measure the performance of students against established international Benchmarks, in the domains of reading, mathematics and science.

If confirmed in the PISA 2021 assessment, these results imply that the performance of Bhutanese students conform, for the vast majority of items, to those that can be expected across current PISA participants of similar proficiency. There is no evidence, in other words, that the test does not capture the full range of proficiency among Bhutan’s students in a way that supports the comparison with international benchmarks.

Overall performance of Bhutan in PISA-D

Let us look at the overall Bhutan performance in PISA-D in respect of its solution rates. By solution rates we mean the average percentage of correct answers on the three tests of reading, mathematics and science. Bhutan has, on average, significantly higher solution rates than the average of all PISA-D countries in all three domains. Table 1 shows the average solution rates (percentage) for Bhutan and all PISA-D countries in Reading literacy. As the table shows, the solution rates for Bhutan are in all three domains in-between those of the two highest-performing PISA-D countries (Ecuador and Honduras), which are therefore selected as reference countries.

<table>
<thead>
<tr>
<th>Domain</th>
<th>BTN</th>
<th>ECU</th>
<th>GTM</th>
<th>HND</th>
<th>KHM</th>
<th>PRY</th>
<th>SEN</th>
<th>ZMB</th>
<th>PISA-D AVG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reading literacy</td>
<td>45.34</td>
<td>52.66</td>
<td>44.50</td>
<td>44.74</td>
<td>34.62</td>
<td>47.03</td>
<td>37.37</td>
<td>27.88</td>
<td>41.26</td>
</tr>
<tr>
<td>Mathematical literacy</td>
<td>38.84</td>
<td>41.69</td>
<td>31.33</td>
<td>31.21</td>
<td>30.29</td>
<td>31.44</td>
<td>27.52</td>
<td>19.08</td>
<td>30.65</td>
</tr>
<tr>
<td>Scientific literacy</td>
<td>45.10</td>
<td>46.98</td>
<td>39.82</td>
<td>40.71</td>
<td>34.85</td>
<td>41.16</td>
<td>33.25</td>
<td>31.19</td>
<td>38.28</td>
</tr>
</tbody>
</table>
However, we recommend that these data be considered with caution as unlike other PISA-D countries, the sample for Bhutan was not designed to provide representative results for the country.

In the remainder of this report we look at the Bhutan performance in PISA-D in more detail in each of the three domains tested: reading, mathematics and science.
Reading literacy assessment in PISA-D

Reading literacy is defined as an individual’s capacity to understand, use, reflect on and engage with written texts, in order to achieve one’s goals, to develop one’s knowledge and potential, and to participate in society.

In reading, the baseline level of skills is defined as the level at which students can not only read simple and familiar texts and understand them literally, but also demonstrate, even in the absence of explicit directions, some ability to connect several pieces of information, draw inferences that go beyond the explicitly stated information, and connect a text to their personal experience and knowledge.

The PISA reading literacy assessment is built on three major task characteristics to ensure a broad coverage of the domain: aspect, text and situation. Figure 1 shows the task characteristics of the reading literacy framework.

Figure 1: The task characteristics of the reading literacy framework (OECD, 2010)
For PISA-D, an additional process titled “literal comprehension” has been added. Literal comprehension requires students to comprehend explicitly stated information that may be found in individual words, sentences or passages. In addition, the concept of “retrieving information” is broadened to range from locating explicitly stated individual pieces of information, such as individual words or phrases, up to finding information in large passages.

Figure 2 shows the relationship between the five processes targeted in the test development for PISA in general and the additional process that was assessed for PISA-D (in orange). The three broad categories reported on subscales in general PISA, when reading literacy is the major domain, are marked in bold. Because there is no major domain in PISA-D, reading literacy will be reported on a single overall scale only.

Solution rates in Reading in Bhutan

Bhutanese students achieved an average solution rate of 45.3 percent in the PISA-D 2017 reading literacy assessment, which was higher than the PISA-D average solution rate of 41.3 percent.

The sections below show the absolute differences between Bhutan’s solution rates and the average solution rates of all PISA-D countries (rates in percentages) and PISA-D reference countries and PISA averages. As mentioned earlier, these data should be considered with caution as they are not nationally representative.

Overall performance and gender differences

Figure 3 visually compare the solution rate of students who took the PISA test in Bhutan to the solution rates typically observed in other countries. Each “dot” in the figure represents a particular item, with the easiest items – those with the highest solution rates – shown in the upper-right corner, and the hardest items closest to the lower-left corner. The blue diagonal line represents a benchmark situation in which students in Bhutan achieve the same solution rate as students in comparison countries on average. The dots in the shaded region are the items in which Bhutan students had lower solution rates than
PISA-D average.

Figure 3: How students in Bhutan fared in PISA compared to the average of all PISA-D countries

Percent-correct scatter plot, Bhutan vs PISA-D average (Reading Literacy)

Figure 3 clearly shows that on the majority of items, students in Bhutan had higher success rates than those in the PISA-D countries. The fact that not all items are perfectly aligned, and that, for a given success rate (and hence, difficulty level) in the comparison countries, there is a relatively wide range of success rates in Bhutan, may reflect both a certain level of statistical uncertainty due to the sampling of items and students, and the existence of peculiar strengths and weaknesses among students in Bhutan. The next section will therefore concentrate, in particular, on the items where the solution rate for Bhutan deviates most significantly from the expected solution rate, given the overall proficiency demonstrated by Bhutanese students in the test. By doing so, we hope to identify patterns in these items and form hypotheses about the peculiar strengths and weaknesses of Bhutan’s students, in comparison with students elsewhere who took part in PISA-D.

A similar comparison of solution rates is also done across subgroups of students. Figure 4 shows how the solution rate of boys who took the PISA-D test in Bhutan compares to the solution rate for girls on the same set of reading items. The red diagonal line represents a benchmark situation in which both boys and girls in Bhutan achieve the same solution
rate. The blue dots represent the average solution rates for each item. In Reading literacy the majority of dots can be seen above the blue line, implying that girls in Bhutan have higher solution rates in reading.

Figure 4: Gender differences in reading performance

Percent-correct scatter plot, Boys vs Girls

Relative strengths

Relative strengths of students in Bhutan are items on which the solution rate is conspicuously higher than expected, given the overall performance of students in Bhutan and the typical patterns observed in other countries.

In order to identify such items, it is necessary to compare the solution rates of all items to their expected level. The percent-correct measure ($P^+$) of solution rates is not well-suited for such comparisons; indeed, the $P^+$ values for very easy items tend to vary less, across countries, than the $P^+$ values for items of average difficulty, for example because a solution rate can never exceed 100%; similarly, the $P^+$ values for very hard items can never be below 0%. An alternative measure of solution rates, so-called delta scores, are better suited to identify the expected level of performance on each item, and to compare success rates across items of varying difficulty. Delta scores are non-linear
transformations of the percent-correct values: The $P^*$ values are converted to z-scores, using an inverse-normal transformation, and then linearly transformed to an expected value of 13.0 and a standard deviation of 4.0. Deltas ordinarily range from 6.0 for a very easy item (approximately 95% correct) to 20.0 for a very hard item (approximately 5% correct), with 13.0 corresponding to 50% correct. The advantage of applying this delta-scale transformation to solution rates is that it “stretches out” differences in solution rates among very easy or very difficult items.

Based on the delta score, and on the reference countries that administered a common set of items, the expected level of performance for Bhutan on each item can be defined as Bhutan’s mean delta score across all common items, plus or minus the difference observed in the reference countries on average between the mean delta score across all common items and the delta score for the item of interest. Conspicuous items are defined as those items where the actual, observed delta score and the expected delta score differ by 1.5 units or more, indicating that the item is significantly harder-than-expected or easier-than-expected.

Table 2: Relative strengths of Bhutan compared to the average of all PISA-D countries

<table>
<thead>
<tr>
<th>ITEM ID</th>
<th>ITEM NAME</th>
<th>BTN</th>
<th>PISA-D</th>
<th>DIFF</th>
<th>ITEM FORMAT</th>
<th>SITUATION</th>
<th>ASPECT</th>
</tr>
</thead>
<tbody>
<tr>
<td>pr460q01</td>
<td>Gulf of Mexico</td>
<td>10.65</td>
<td>13.80</td>
<td>-3.15</td>
<td>Open response</td>
<td>Educational</td>
<td>Access and retrieve</td>
</tr>
<tr>
<td>pr442q02</td>
<td>Galileo</td>
<td>9.39</td>
<td>12.02</td>
<td>-2.63</td>
<td>Open response</td>
<td>Personal</td>
<td>Access and retrieve</td>
</tr>
<tr>
<td>pr455q02</td>
<td>Chocolate and Health</td>
<td>15.76</td>
<td>17.98</td>
<td>-2.22</td>
<td>Open response</td>
<td>Personal</td>
<td>Access and retrieve</td>
</tr>
<tr>
<td>pr6018q3a</td>
<td>Aski Gym Youth Programme</td>
<td>15.94</td>
<td>18.08</td>
<td>-2.14</td>
<td>Open response</td>
<td>Personal</td>
<td>Integrate and interpret</td>
</tr>
<tr>
<td>pr456q01s</td>
<td>Biscuits</td>
<td>5.95</td>
<td>8.09</td>
<td>-2.14</td>
<td>Multiple choice</td>
<td>Personal</td>
<td>Access and retrieve</td>
</tr>
</tbody>
</table>

This method reveals 5 conspicuous items with significant differences in the delta scores of Bhutan compared to the average of all the PISA-D countries. It is noticeable that the 4 out of 5 items are relatively easy belonging to “Access and retrieve” aspect. These items do not require students to use outside information, and require only minimal inferences.

That these items constitute relative strengths for Bhutan can be seen even more simply when we compare solution rates (as measured by “percent correct) in Bhutan with the percent correct of these items in the five reference countries and PISA average.

Table 3: Relative strengths of Bhutan compared to the reference countries

<table>
<thead>
<tr>
<th>ITEM ID</th>
<th>ITEM NAME</th>
<th>BTN</th>
<th>ECU</th>
<th>PRY</th>
<th>PISA-D AVG</th>
<th>SGP</th>
<th>SWE</th>
<th>DOM</th>
<th>OECD AVG</th>
</tr>
</thead>
<tbody>
<tr>
<td>pr460q01</td>
<td>Gulf of Mexico</td>
<td>72.16</td>
<td>53.74</td>
<td>53.17</td>
<td>42.23</td>
<td></td>
<td></td>
<td></td>
<td>NA</td>
</tr>
<tr>
<td>pr442q02</td>
<td>Galileo</td>
<td>81.67</td>
<td>77.73</td>
<td>72.97</td>
<td>59.19</td>
<td></td>
<td></td>
<td></td>
<td>NA</td>
</tr>
<tr>
<td>pr455q02</td>
<td>Chocolate and Health</td>
<td>24.47</td>
<td>23.84</td>
<td>17.61</td>
<td>12.99</td>
<td></td>
<td></td>
<td></td>
<td>NA</td>
</tr>
<tr>
<td>pr6018q3a</td>
<td>Aski Gym Youth Programme</td>
<td>23.08</td>
<td>15.45</td>
<td>16.08</td>
<td>11.25</td>
<td></td>
<td></td>
<td></td>
<td>NA</td>
</tr>
<tr>
<td>pr456q01s</td>
<td>Biscuits</td>
<td>96.10</td>
<td>96.02</td>
<td>92.61</td>
<td>87.48</td>
<td>97.57</td>
<td>92.63</td>
<td>90.82</td>
<td>95.35</td>
</tr>
</tbody>
</table>

Students in Bhutan have higher solution rates in open response compared to multiple choice items. They also had better performance in items related to personal situations. This observation is also further confirmed by aggregated delta values for the items formats and situations. Table 7 shows that, of the three aspects, Bhutanese students do best on access and retrieve task.
Relative weaknesses

Table 4: Relative weaknesses of Bhutan compared to the average of all PISA-D countries

<table>
<thead>
<tr>
<th>ITEM ID</th>
<th>ITEM NAME</th>
<th>BTN</th>
<th>PISA-D</th>
<th>DIFF</th>
<th>ITEM FORMAT</th>
<th>SITUATION</th>
<th>ASPECT</th>
</tr>
</thead>
<tbody>
<tr>
<td>pr432q01</td>
<td>About a book</td>
<td>14.32</td>
<td>13.24</td>
<td>1.08</td>
<td>Open response</td>
<td>Personal</td>
<td>Integrate and interpret</td>
</tr>
<tr>
<td>pr442q05</td>
<td>Galileo</td>
<td>10.73</td>
<td>18.56</td>
<td>7.83</td>
<td>Open response</td>
<td>Personal</td>
<td>Reflect and evaluate</td>
</tr>
<tr>
<td>pr101q03s</td>
<td>Rhino</td>
<td>16.15</td>
<td>14.60</td>
<td>1.55</td>
<td>MCQ</td>
<td>Public</td>
<td>Reflect and evaluate</td>
</tr>
<tr>
<td>pr101q02s</td>
<td>Rhino</td>
<td>14.39</td>
<td>11.47</td>
<td>2.92</td>
<td>MCQ</td>
<td>Public</td>
<td>Integrate and interpret</td>
</tr>
<tr>
<td>pr101q01s</td>
<td>Rhino</td>
<td>18.03</td>
<td>14.90</td>
<td>3.13</td>
<td>MCQ</td>
<td>Public</td>
<td>Integrate and interpret</td>
</tr>
</tbody>
</table>

Table 4 displays five conspicuous items relating to positive differences between the delta values of Bhutan and those observed on average of across all PISA-D countries. It is noticeable that no access and retrieve item is found among these items – all items on which Bhutan’s students performed below the expected level, involve either “integrate and interpret” or “reflect and evaluate”.

The weakness of Bhutanese students can be seen in items related to public and personal situation. Students found items which require forming broad understanding and developing an interpretation of a text, and reflecting on and evaluating content or form of a text difficult.

To classify the performance of Bhutan in more detail, the items identified in Table 5 are taken and we compare percent correct of Bhutan with the percent correct of these items in the five reference countries, PISA-D and OECD averages.

Table 5: Relative weaknesses of Bhutan compared to the reference countries

<table>
<thead>
<tr>
<th>ITEM ID</th>
<th>ITEM NAME</th>
<th>BTN</th>
<th>ECU</th>
<th>PRY</th>
<th>PISA-D AVG</th>
<th>SGP</th>
<th>SWE</th>
<th>DOM</th>
<th>OECD AVG</th>
</tr>
</thead>
<tbody>
<tr>
<td>pr432q01</td>
<td>About a book</td>
<td>37.05</td>
<td>56.01</td>
<td>51.13</td>
<td>47.91</td>
<td>NA</td>
<td></td>
<td></td>
<td>64.58</td>
</tr>
<tr>
<td>pr101q03s</td>
<td>Rhino</td>
<td>21.56</td>
<td>51.75</td>
<td>43.44</td>
<td>35.18</td>
<td>65.05</td>
<td>70.48</td>
<td>NA</td>
<td>84.59</td>
</tr>
<tr>
<td>pr442q05</td>
<td>Galileo</td>
<td>4.63</td>
<td>13.88</td>
<td>18.07</td>
<td>10.35</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pr101q02s</td>
<td>Rhino</td>
<td>36.45</td>
<td>77.71</td>
<td>75.59</td>
<td>64.26</td>
<td>80.90</td>
<td>82.86</td>
<td>NA</td>
<td>84.59</td>
</tr>
<tr>
<td>pr101q01s</td>
<td>Rhino</td>
<td>10.42</td>
<td>41.19</td>
<td>42.80</td>
<td>32.09</td>
<td>54.10</td>
<td>49.21</td>
<td>NA</td>
<td>52.55</td>
</tr>
</tbody>
</table>

It is interesting to see that Bhutanese students’ strengths mostly manifest on easier items, for example, items in strengths had delta values from 5.95 to 15.94. And weakness patterns mostly on difficult items, for example, items in weakness had delta values from 14.32 to 19.73. It is clear that Bhutan performed better in easy items and rather moderately in difficult items.

Although Bhutan performed better than most PISA-D countries, there is a huge performance gap between Bhutan and PISA countries. Bhutan’s performance is about 38 to 48 points (percentage points) below the percent correct values of PISA reference countries.

Table 6: Average deltas for reading item classification of PISA-D countries

<table>
<thead>
<tr>
<th>ITEM GROUP</th>
<th>BTN</th>
<th>ECU</th>
<th>GTM</th>
<th>HND</th>
<th>KHM</th>
<th>PRY</th>
<th>SEN</th>
<th>ZMB</th>
<th>PISA-D AVG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Access and retrieve</td>
<td>11.21</td>
<td>10.51</td>
<td>11.49</td>
<td>11.41</td>
<td>12.59</td>
<td>11.32</td>
<td>12.80</td>
<td>14.07</td>
<td>12.03</td>
</tr>
<tr>
<td>Reflect and evaluate</td>
<td>15.29</td>
<td>14.10</td>
<td>15.38</td>
<td>15.60</td>
<td>16.87</td>
<td>15.22</td>
<td>16.12</td>
<td>18.07</td>
<td>15.91</td>
</tr>
</tbody>
</table>

...
Analysis of illustrating items

In this section, we will illustrate some of the patterns observed in Table 2 and 4 with a typical PISA release item. Each item in the PISA test has its own properties and its specific cognitive requirements.

A. Strength patterns

To illustrate the pattern, we have chosen item “Miser” and “The Galapagoes Islands”. These items’ characteristics are similar to the strengths pattern observed for Bhutan.

**Item 1**  The Miser and his Gold

A fable by Aesop

A miser sold all that he had and bought a lump of gold, which he buried in a hole in the ground by the side of an old wall. He went to look at it daily. One of his workmen observed the miser’s frequent visits to the spot and decided to watch his movements. The workman soon discovered the secret of the hidden treasure, and digging down, came to the lump of gold, and stole it. The miser, on his next visit, found the hole empty and began to tear his hair and to make loud lamentations. A neighbour, seeing him overcome with grief and learning the cause, said, “Pray do not grieve so; but go and take a stone, and place it in the hole, and fancy that the gold is still lying there. It will do you quite the same service; for when the gold was there, you had it not, as you did not make the slightest use of it.”

Use the fable “The Miser and his Gold” on the previous page to answer the questions that follow.

How did the miser get a lump of gold?

Source: OECD 2014 What Students Know and Can Do: Student Performance in Mathematics, Reading and Science – Volume I.

This is one of the easiest questions in PISA reading, where student is required to access and retrieve a piece of explicitly stated information. In order to answer the question, student has to either quote directly from the text - “He sold all that he had” – or provide a paraphrase such as “He sold all his stuff”. The formal language of the text, which is likely to have added difficulty in other questions in the unit, is unlikely to have much impact here because the required information is located at the very beginning of the text. Although this is an easy question in PISA’s frame of reference, it still requires a small degree of inference, beyond the absolutely literal: the student must infer that there is a causal connection between the first proposition (that the miser sold all he had) and the second (that he bought gold).
According to the Conservation webpage, what was the main goal for why conservationists started a breeding program for tortoises?

- To save the tortoises from extinction.
- To monitor how tortoises mature.
- To protect tortoise eggs from predators.
- To track tortoises for a long period of time.

The Galapagos Islands

Located 1000 kilometres west of the South American coast lie the Galapagos Islands – one of the most fascinating places in the world.

There are currently 95 indigenous species of animals that exist solely on the various islands of the archipelago. Many people travel to the Galapagos Islands to observe these special animals in their natural habitat. The islands are often referred to as a “living laboratory” because they offer scientists great research potential. Being near the equator, the islands receive ample sunshine, while the strong ocean currents provide cool breezes. Many plants and animals thrive in this environment. Tourists and scientists alike are fascinated by the animals who seem just as curious about humans as we are about them. Galapagos animals evolved for centuries without human interference or predation and consequently, when approached by humans, they don’t show fear like most animals throughout the world. They often wander up to visitors! This behavior creates amazing photo opportunities, but it has made the animals very vulnerable.

Over the years, the ecosystem surrounding the Galapagos Islands has been threatened due to human activity on the islands. Damage to the ecosystem has had negative consequences on populations of many of the Galapagos animals. Thankfully, with the work of committed researchers, the ecosystem is slowly recovering.

This item do not explicitly identify the relevant text for the student. Thus, searching for the relevant text is required to complete the item accurately. Here, the item explicitly refers to the “Conservation” webpage. This should be a strong signal to the student that they need to navigate to the “Conservation” webpage to find the answer. Once students are on the correct webpage, they need to match the information in the question stem (started a breeding program for tortoises) and the webpage (launched a breeding program). Option A is also a very close match with what is in the webpage (...to save the rest of the tortoises from extinction). The correct answer is (A) To save the tortoises from extinction.

By contrasting this item and the previous two items, one can see the difference between the two cognitive processes of access and retrieve information within a text, and search for and select relevant text.

In PISA-D, which was administered as a paper-based test, the process “search for and select relevant text” (among multiple texts) was not evaluated. As an important process in online reading, this process features prominently in the digital version of the PISA test, and will be assessed in PISA 2021.
B. Weakness patterns

To illustrate the pattern, we have chosen “Supermarket Notice” and “Tall Buildings”. These items’ characteristics are similar to the weakness patterns observed for Bhutan.

Item 1  Supermarket Notice

---

**Peanut Allergy Alert**

**Lemon Cream Biscuits**

**Date of alert**: 04 February  
**Manufacturer’s Name**: Fine Foods Ltd  
**Product Information**: 125g Lemon Cream Biscuits (Best before 18 June and Best before 01 July)  
**Details**: Some biscuits in these batches may contain pieces of peanut, which are not included in the ingredient list. People with an allergy to peanuts should not eat these biscuits.  
**Consumer action**: If you have bought these biscuits you may return the product to the place of purchase for a full refund.  
**Or call 1800 034 241 for further information**

Why does the notice include “Best before” dates?

Ans: __________________________________________________________________________

Source: PISA 2009 Assessment Framework – Key Competencies in Reading, Mathematics and Science

---

<table>
<thead>
<tr>
<th>Situation</th>
<th>Public</th>
</tr>
</thead>
<tbody>
<tr>
<td>Medium</td>
<td>Print</td>
</tr>
<tr>
<td>Text format</td>
<td>Non-continuous</td>
</tr>
<tr>
<td>Text type</td>
<td>Instruction</td>
</tr>
<tr>
<td>Aspect</td>
<td>Integrate and interpret: Develop an interpretation</td>
</tr>
<tr>
<td>Question intent</td>
<td>Identify the purpose of a conventional feature included in a short text</td>
</tr>
<tr>
<td>Item format</td>
<td>Open constructed response</td>
</tr>
</tbody>
</table>

This question was answered correctly by less than one-third of students across OECD countries. Given the shortness and simplicity of the text, this illustrates the fact that the characteristics of a text only partly explain the difficulty of an item. The question requires students to identify the purpose of a specified part of the text, namely, the “best before dates”. The difficulty of the item comes from the fact that students must focus on the
purpose of the feature in this particular text. Students who answer by giving the usual purpose of this feature (that is, to tell the consumer when the product should be used by) do not receive credit for this item. In this respect the full credit response is contrary to expectations, an established marker of item difficulty.

**Item 2 Tall Buildings**

“Tall buildings” is an article from a Norwegian magazine published in 2006 ...

**Figure 1: Tall buildings of the world**

Figure 1 shows the number of buildings of at least 30 storeys that have been built, or are under construction. This includes buildings that have been proposed since January 2001.

**Figure 2: Some of the world’s tallest buildings**

The Burj Tower in Dubai is expected to be the tallest building in the world, at 700 metres, when it is finished in 2008.

The Radisson SAS Plaza in Oslo, Norway is only 117 metres tall. Why has it been included in Figure 2?

Source: PISA 2009 Assessment Framework – Key Competencies in Reading, Mathematics and Science

**Situation**

Educational

**Medium**

Print

**Text format**

Non-continuous

**Text type**

Exposition

**Aspect**

Reflect and evaluate: Reflect on and evaluate the content of a text

**Question intent**

Recognise the influence of reader’s perspective on the way a text is constructed

**Item format**

Open constructed response

This item was difficult. Only about one-quarter of students answered correctly. Students
are require to recognise that the purpose for including a specified building in Figure 2. Which is achieved by relating the information in the introduction (that the article was published in a Norwegian magazine) to the author's decision to include the tallest building in Norway (the Radisson SAS Plaza) in Figure 2 or by expressing the idea of perspective in general, rather than in specific terms.
Mathematical literacy assessment in PISA-D

Mathematical literacy is defined as an individual’s capacity to formulate, employ, and interpret mathematics in a variety of contexts. It includes reasoning mathematically and using mathematical concepts, procedures, facts and tools to describe, explain and predict phenomena. It assists individuals to recognise the role that mathematics plays in the world and to make the well-founded judgments and decisions needed by constructive, engaged and reflective citizens.

In mathematics, the baseline level of skills is defined as the level at which students can not only carry out routine procedures, such as an arithmetic operation, in situations where all the instructions are given to them, but can also interpret and recognise how a (simple) situation (e.g. comparing the total distance across two alternative routes, or converting prices into a different currency) can be represented mathematically.

In short, mathematical literacy describes one’s capacity to use mathematics in a well-founded manner in order to solve real world problems, where “real world” means “…the physical, social and mental world…” (Freudenthal, 1983), including the mathematical world itself. Figure 5 (OECD, 2013a: 26) shows the processes involved when solving real world problems by means of mathematics.

Figure 5: A model of mathematical literacy in practice

The processes “Formulate – Employ – Interpret – Evaluate” form the so-called modelling cycle. Underlying these processes are certain competencies (Niss and Hojgaard, 2011) called “fundamental mathematical capabilities” in the PISA framework. Seven such competencies are distinguished:

- Communicating
- Representing
- Devising strategies
- Mathematising
- Reasoning and arguing
- Using symbolic and formal language
- Using tools

The mathematical contents are organised into four categories: 1) quantity; 2) change and relationships; 3) space and shape; and 4) uncertainty and data. These categories are not to be confused with the traditional mathematical content areas for 15-year-olds. However, there are certain connections that are stronger than others: quantity with arithmetics; change and relationships with algebra; space and shape with geometry; and uncertainty and data with probability and statistics.

Solution rates in Mathematics in Bhutan

Bhutanese students achieved an average solution rate of 38.8 percent in the PISA-D 2017 mathematical literacy assessment, which was significantly higher than the PISA-D average solution rate of 30.6 percent.

The sections below show the absolute differences between Bhutan’s solution rates in Mathematics and the average solution rates of all PISA-D countries (rates in percentages) and the PISA-D reference countries.

Overall performance and gender differences

Figure 6 visually compared the solution rate of students who took the PISA Mathematics test in Bhutan to the solution rates typically observed in other countries. Each “dot” in the figure represents a particular item, with the easiest items – those with the highest solution rates – shown in the upper-right corner, and the hardest items closest to the lower-left corner. The diagonal line represents a benchmark situation in which students in Bhutan achieve the same solution rate as students in comparison countries on average. The dots in the shaded region are the items in which Bhutan students had lower solution rates than PISA-D average.
Figure 6: How students in Bhutan fared in PISA compared to the average of all PISA-D countries

Percent-correct scatter plot, Bhutan vs PISA-D average (Mathematical Literacy)

The visual inspection of Figures 6 clearly shows that on the majority of items, students in Bhutan had higher success rates. The fact that not all items are perfectly aligned, and that, for a given success rate (and hence, difficulty level) in the comparison countries, there is a relatively wide range of success rates in Bhutan, may reflect both a certain level of statistical uncertainty due to the sampling of items and students, and the existence of peculiar strengths and weaknesses among students in Bhutan. The next section will therefore concentrate, in particular, on the items where the solution rate for Bhutan deviates most significantly from the expected solution rate, given the overall proficiency demonstrated by Bhutanese students in the test.

A similar comparison of solution rates can also be done across subgroups of students. Figure 7 shows how the solution rate of boys who took the PISA test in Bhutan compares to the solution rate for girls on the same set of Mathematics items. The blue dots represent the average solution rates for each item. In Mathematical literacy majority of dots can be seen below the blue line, implying that boys in Bhutan have higher solution rates in mathematical literacy.
Figure 7: Gender differences in mathematics performance

Percent-correct scatter plot, Boys vs Girls

Relative strengths

As in reading, relative strengths and weaknesses of students in Bhutan can be identified by applying a non-linear transformation to the solution rate on each item, then comparing the difference between this so-called “delta-score” and the average delta across all items with the corresponding difference in comparison countries.

Table 7: Relative strengths of Bhutan compared to the average of all PISA-D countries

<table>
<thead>
<tr>
<th>ITEM ID</th>
<th>ITEM NAME</th>
<th>DELTA</th>
<th>DIFF</th>
<th>SITUATION / CONTEXT</th>
<th>CONTENT</th>
<th>PROCESS</th>
</tr>
</thead>
<tbody>
<tr>
<td>pm915q02</td>
<td>Carbon Tax</td>
<td>12.2</td>
<td>15.6</td>
<td>-2.5</td>
<td>Societal</td>
<td>Change and Relationships</td>
</tr>
<tr>
<td>pm955q01</td>
<td>Migration</td>
<td>12.5</td>
<td>15.4</td>
<td>-1.9</td>
<td>Societal</td>
<td>Uncertainty and Data</td>
</tr>
<tr>
<td>pm5169q01</td>
<td>Shoe Sizes</td>
<td>13.5</td>
<td>16.3</td>
<td>-1.9</td>
<td>Personal</td>
<td>Change and Relationships</td>
</tr>
<tr>
<td>pm936q01</td>
<td>Seats in a Theatre</td>
<td>14.3</td>
<td>17.0</td>
<td>-1.8</td>
<td>Occupational</td>
<td>Change and Relationships</td>
</tr>
</tbody>
</table>
This method reveals 7 conspicuous items with significant differences in the delta scores of Bhutan compared to the average of all the PISA-D countries. Four of these items require employing mathematical concepts, facts and procedures; three require interpreting, applying and evaluating mathematical outcomes. They are mostly from the mathematical content area “Change and relationship” or from “Uncertainty and data”. Almost all of the conspicuous items are from societal context.

To classify the performance of Bhutan in more detail, the items identified in Table 9 are taken and we compare percent correct of Bhutan with the percent correct of these items in PISA-D and PISA reference countries.

Table 8: Relative strengths of Bhutan compared to the reference countries

<table>
<thead>
<tr>
<th>ITEM ID</th>
<th>ITEM NAME</th>
<th>BTN</th>
<th>ECU</th>
<th>PRY</th>
<th>PISA-D AVG</th>
<th>SGP</th>
<th>SWE</th>
<th>DOM</th>
<th>OECD AVG</th>
</tr>
</thead>
<tbody>
<tr>
<td>pm915q02</td>
<td>Carbon Tax</td>
<td>57.63</td>
<td>27.52</td>
<td>20.04</td>
<td>25.47</td>
<td>85.71</td>
<td>50.12</td>
<td>8.66</td>
<td>64.17</td>
</tr>
<tr>
<td>pm955q01</td>
<td>Migration</td>
<td>55.07</td>
<td>42.09</td>
<td>26.01</td>
<td>27.84</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>pm5169q01</td>
<td>Shoe Sizes</td>
<td>45.35</td>
<td>36.91</td>
<td>11.22</td>
<td>20.57</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>pm936q01</td>
<td>Seats in a Theatre</td>
<td>37.22</td>
<td>27.08</td>
<td>7.21</td>
<td>15.58</td>
<td>NA</td>
<td>NA</td>
<td>8.71</td>
<td>30.78</td>
</tr>
<tr>
<td>pm982q01</td>
<td>Employment Data</td>
<td>82.70</td>
<td>72.71</td>
<td>64.66</td>
<td>61.99</td>
<td>87.00</td>
<td>86.07</td>
<td>51.98</td>
<td>84.97</td>
</tr>
<tr>
<td>pm650p001</td>
<td>Urban Population</td>
<td>50.47</td>
<td>41.34</td>
<td>28.04</td>
<td>27.23</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
</tbody>
</table>

On their strongest items, students in Bhutan performed significantly better compared to top PISA-D countries but slightly (2 to 4%) below students in Sweden and significantly below students in Singapore.

Relative weaknesses

Table 9: Relative weaknesses of Bhutan compared to the average of all PISA-D countries

<table>
<thead>
<tr>
<th>ITEM ID</th>
<th>ITEM NAME</th>
<th>DELTA</th>
<th>DIFF</th>
<th>ITEM FORMAT</th>
<th>CONTENT</th>
<th>PROCESS</th>
</tr>
</thead>
<tbody>
<tr>
<td>pm496q02</td>
<td>Cash Withdrawal</td>
<td>16.6</td>
<td>3.1</td>
<td>Open response</td>
<td>Quantity</td>
<td>Employing Mathematical Concepts, Facts and Procedures</td>
</tr>
<tr>
<td>pm919q02</td>
<td>Fan Merchandise</td>
<td>19.3</td>
<td>2.4</td>
<td>Open response</td>
<td>Quantity</td>
<td>Formulating Situations Mathematically</td>
</tr>
<tr>
<td>pm909q03</td>
<td>Speeding Fines</td>
<td>20.7</td>
<td>2.0</td>
<td>Open response</td>
<td>Change and Relationships</td>
<td>Interpreting, Applying and Evaluating Mathematical Outcomes</td>
</tr>
<tr>
<td>pm604p505a</td>
<td>Gas Gauge</td>
<td>15.3</td>
<td>1.8</td>
<td>Human coded</td>
<td>Quantity</td>
<td>Employing Mathematical Concepts, Facts and Procedures</td>
</tr>
<tr>
<td>pm954q02</td>
<td>Medicine Doses</td>
<td>20.3</td>
<td>1.7</td>
<td>Open response</td>
<td>Change and Relationships</td>
<td>Employing Mathematical Concepts, Facts and Procedures</td>
</tr>
</tbody>
</table>
Table 9 displays 8 conspicuous items relating to differences between the delta values of Bhutan and the average of all PISA-D countries. Three of these weak items require employing mathematical concepts, facts and procedures; three require formulating situations mathematically; and two require interpreting, applying and evaluating mathematical outcomes. Four of the relative weaknesses are from the content area “Quantity”.

To classify the performance of Bhutan in more detail, the items identified in Table 11 are taken and we compare percent correct of Bhutan with the percent correct of these items in PISA-D and PISA reference countries.

Table 10: Relative weaknesses of Bhutan compared to the reference countries

<table>
<thead>
<tr>
<th>ITEM ID</th>
<th>ITEM NAME</th>
<th>BTN</th>
<th>ECU</th>
<th>PRY</th>
<th>PISA-D AVG</th>
<th>SGP</th>
<th>SWE</th>
<th>DOM</th>
<th>OECD AVG</th>
</tr>
</thead>
<tbody>
<tr>
<td>pm496q02</td>
<td>Cash Withdrawal</td>
<td>18.70</td>
<td>56.41</td>
<td>45.25</td>
<td>36.40</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pm919q02</td>
<td>Fan Merchandise</td>
<td>5.71</td>
<td>22.52</td>
<td>12.68</td>
<td>11.00</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pm909q03</td>
<td>Speeding Fines</td>
<td>2.70</td>
<td>7.52</td>
<td>4.57</td>
<td>4.71</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pm604p001</td>
<td>Gas Gauge</td>
<td>28.23</td>
<td>52.29</td>
<td>44.03</td>
<td>35.26</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pm954q02</td>
<td>Medicine Doses</td>
<td>3.48</td>
<td>12.27</td>
<td>6.34</td>
<td>5.27</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pm446q01</td>
<td>Thermometer Cricket</td>
<td>22.82</td>
<td>44.33</td>
<td>39.44</td>
<td>28.76</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pm510q01s</td>
<td>Baby Growth</td>
<td>16.97</td>
<td>38.92</td>
<td>23.39</td>
<td>20.56</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pm645p001</td>
<td>Airport Timetable</td>
<td>16.38</td>
<td>32.42</td>
<td>25.16</td>
<td>19.74</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The contrast between strengths and weaknesses in Table 7 and 9 suggests that students in Bhutan may find it particularly difficult to formulate situations mathematically, and with tasks related to the quantity content area.

It is interesting to see that delta values of items in strengths area ranged from 9.2 to 14.3 and delta values in weakness area ranged from 15.3 to 20.7. It is clear that Bhutanese students found mathematics questions generally difficult.

Table 11: Average deltas for mathematics item classification of PISA-D countries

<table>
<thead>
<tr>
<th>ITEM GROUP</th>
<th>BTN</th>
<th>ECU</th>
<th>GTM</th>
<th>HND</th>
<th>KHM</th>
<th>PRY</th>
<th>SEN</th>
<th>ZMB</th>
<th>PISA-D AVG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Change and relationships</td>
<td>15.51</td>
<td>15.00</td>
<td>15.03</td>
<td>16.24</td>
<td>16.43</td>
<td>16.92</td>
<td>16.67</td>
<td>18.54</td>
<td>16.70</td>
</tr>
<tr>
<td>Quantity</td>
<td>14.47</td>
<td>13.81</td>
<td>15.17</td>
<td>14.89</td>
<td>15.53</td>
<td>15.31</td>
<td>15.59</td>
<td>17.39</td>
<td>15.38</td>
</tr>
<tr>
<td>Space and shape</td>
<td>13.98</td>
<td>13.71</td>
<td>14.95</td>
<td>14.88</td>
<td>14.82</td>
<td>15.07</td>
<td>15.01</td>
<td>15.30</td>
<td>16.64</td>
</tr>
<tr>
<td>Uncertainty and data</td>
<td>14.02</td>
<td>13.96</td>
<td>15.07</td>
<td>15.91</td>
<td>15.44</td>
<td>14.93</td>
<td>15.49</td>
<td>16.52</td>
<td>15.20</td>
</tr>
<tr>
<td>Scientific</td>
<td>15.21</td>
<td>14.41</td>
<td>15.92</td>
<td>15.91</td>
<td>16.05</td>
<td>15.60</td>
<td>16.22</td>
<td>18.30</td>
<td>16.06</td>
</tr>
<tr>
<td>Societal</td>
<td>14.63</td>
<td>14.39</td>
<td>15.46</td>
<td>15.23</td>
<td>15.86</td>
<td>15.43</td>
<td>16.00</td>
<td>17.24</td>
<td>15.66</td>
</tr>
</tbody>
</table>
Analysis of illustrating items

In this section, we will illustrate some of the patterns observed in Table 7 and 9 with a typical PISA release item. Each item in the PISA test has its own properties and its specific cognitive requirements.

A. Strength patterns

To illustrate the pattern, we have chosen item “Helen The Cyclist” and “Charts”. These items’ characteristics are similar to the strengths patterns observed for Bhutan.

**Item 1: Helen the Cyclist**

Helen has just got a new bike. It has a speedometer which sits on the handlebar.

The speedometer can tell Helen the distance she travels and her average speed for a trip.

**Question (Level 2)**

On one trip, Helen rode 4 km in the first 10 minutes and then 2 km in the next 5 minutes.

Which one of the following statements is correct?

A. Helen’s average speed was greater in the first 10 minutes than in the next 5 minutes.

B. Helen’s average speed was the same in the first 10 minutes and in the next 5 minutes.

C. Helen’s average speed was less in the first 10 minutes than in the next 5 minutes.

D. It is not possible to tell anything about Helen’s average speed from the information given.

Ans: ________________________________

Source: OECD 2014 What Students Know and Can Do: Student Performance in Mathematics, Reading and Science – Volume I

This is a simple multiple-choice item, requires comparison of speed when travelling 4 km in 10 minutes versus 2 km in 5 minutes. It is been classified within the employing process category because it requires the precise mathematical understanding that speed is a rate and that proportionality is the key. This question can be solved by recognising the doubles involved (2 km – 4 km; 5 km – 10 km), which is the very simplest notion of
proportion. Consequently, successful students demonstrate a very basic understanding of speed and of proportion calculations. If distance and time are in the same proportion, the speed is the same. Of course, students could correctly solve the problem in more complicated ways (e.g. calculating that both speeds are 24 km per hour) but this is not necessary. The correct response option here is B (Helen’s average speed was the same in the first 10 minutes and in the next 5 minutes).

<table>
<thead>
<tr>
<th>Content</th>
<th>Change and relationship</th>
</tr>
</thead>
<tbody>
<tr>
<td>Process</td>
<td>Employing Mathematical Concepts, Facts and Procedures</td>
</tr>
<tr>
<td>Item format</td>
<td>Simple multiple choice</td>
</tr>
</tbody>
</table>

**Item 2 - Charts**

In January, the new CDs of the bands 4U2Rock and The Kicking Kangaroos were released. In February, the CDs of the bands No One’s Darling and The Metalfolkies followed. The following graph shows the sales of the bands’ CDs from January to June.

**Question (Level 1)**

How many CDs did the band The Metalfolkies sell in April?

A. 250  
B. 500  
C. 1000  
D. 1270  

Ans: ______________________________________

The task “Charts” is embedded in a real world situation and starts with a short stimulus that explains the structure of the following diagram. The number of CDs sold per month from four different bands is given for the six months from January to June. For each month from February on, there are four bars. Each bar represents the number of CDs sold in a specific month by a single band. The task for the students is to find out how many CDs the band The Metalfolkies sold in April. Therefore, one needs to have a look at the bright grey bar for April, which represents the band The Metalfolkies. It reaches up to 500 so answer B is correct.

B. Weakness patterns

To illustrate the pattern, we have chosen item “Climbing Mount Fuji” and “Helen the Cyclist”. These items’ characteristics are similar to the weakness patterns observed for Bhutan.

Item 1  Climbing Mount Fuji

Mount Fuji is a famous dormant volcano in Japan.

Question (Level 5)

Mount Fuji is only open to the public for climbing from 1 July to 27 August each year. About 200 000 people climb Mount Fuji during this time.

On average, about how many people climb Mount Fuji each day?

A. 340  
B. 710  
C. 3,400  
D. 7,100  
E. 7,400

Ans: ___________________________________

Source: OECD 2014 What Students Know and Can Do: Student Performance in Mathematics, Reading and Science – Volume I
This task was allocated to the formulating category because most of the cognitive effort in this relatively easy item requires taking two pieces of real-world information (open season and total number of climbers) and establishing a mathematical problem to be solved: find the length of the open season from the dates and use it with the information about the total number of climbers to find the average number of climbers each day. Expert judgement is that the major cognitive demand for 15-year-olds lies in this movement from the real world problem to the mathematical relationships, rather than in the ensuing whole number calculations.

**Item 2  Helen the Cyclist**

Helen has just got a new bike. It has a speedometer which sits on the handlebar.
The speedometer can tell Helen the distance she travels and her average speed for a trip.

**Question (Level 6)**

Helen rode her bike from home to the river, which is 4 km away. It took her 9 minutes. She rode home using a shorter route of 3 km. This only took her 6 minutes.

What was Helen's average speed, in km/h, for the trip to the river and back?

Average speed for the trip: ________ km/h
Scientific literacy assessment in PISA-D

Scientific literacy is defined as the ability to engage with science-related issues, and with the ideas of science, as a reflective citizen. A scientifically literate person is willing to engage in reasoned discourse about science and technology which requires the competencies to explain phenomena scientifically, evaluate and design scientific enquiry, and interpret data and evidence scientifically.

In science, the baseline level of proficiency corresponds to the level at which students can draw on their knowledge of basic science content and procedures to interpret data, identify the question being addressed in a simple experiment, or identify whether a conclusion is valid based on the data provided.

The characteristics of the science domain assessed are shown in Figure 8.

Figure 8: Aspects of the scientific literacy assessment framework

- **Contexts**
  - Issues:
    - Personal
    - Local / National
    - Global both current and historical, which demand some understanding of science and technology

- **Scientific competencies**
  - The ability to:
    - Explain phenomena scientifically
    - Evaluate and design scientific enquiry
    - Interpret data and evidence scientifically

- **Scientific knowledge**
  - An understanding of the major facts, concepts and explanatory theories that form the basis of scientific knowledge
  - The three forms of scientific knowledge are:
    - Content knowledge (knowledge of both the natural world and technological artefacts)
    - Procedural knowledge* (knowledge of how such ideas are produced)
    - Epistemic knowledge* (understanding of the underlying rationale for these procedures and the justification for their use)

- **Attitudes towards science**
  - A set of attitudes towards science indicated by:
    - an interest in science and technology
    - valuing scientific approaches to enquiry where appropriate
    - perceiving and being aware of environmental issues

*Procedural knowledge and epistemic knowledge are considered as important aspects of scientific literacy.
*Although different under the theoretical point of view, the procedural and epistemological knowledge categories are part of a single reference category.

Solution rates in Science in Bhutan

Bhutanese students achieved an average solution rate of 45.1 percent in the PISA-D 2017 scientific literacy assessment, which was significantly higher than the PISA-D average solution rate of 38.3 percent.

The sections below show the absolute differences between Bhutan’s solution rates and the average solution rates of all PISA-D countries (rates in percentages) and the reference countries.

Overall performance and gender differences

Figure 9 visually compares the solution rate of students who took the PISA Science test in Bhutan to the solution rates typically observed in other countries. Each “dot” in the figure represents a particular item, with the easiest items – those with the highest solution rates – shown in the upper-right corner, and the hardest items closest to the lower-left corner. The diagonal line represents a benchmark situation in which students in Bhutan achieve the same solution rate as students in comparison countries on average. The dots in the shaded region are the items in which Bhutan students had lower solution rates than PISA-D average.

Figure 9: How students in Bhutan fared in PISA compared to the average of all PISA-D countries
The visual inspection of Figure 9 clearly shows that on the majority items, students in Bhutan had higher success rates compared to PISA-D averages. The fact that not all items are perfectly aligned, and that, for a given success rate (and hence, difficulty level) in the comparison countries, there is a relatively wide range of success rates in Bhutan, may reflect both a certain level of statistical uncertainty due to the sampling of items and students, and the existence of peculiar strengths and weaknesses among students in Bhutan. The next section will therefore concentrate, in particular, on the items where the solution rate for Bhutan deviates most significantly from the expected solution rate, given the overall proficiency demonstrated by Bhutanese students in the test.

A similar comparison of solution rates can also be done across subgroups of students. Figure 10 shows how the solution rate of boys who took the PISA test in Bhutan compares to the solution rate for girls on the same set of science items. The blue dots represent the average solution rates for each item. In scientific literacy, the number of dots are slightly more below the blue line, implying that boys in Bhutan have slightly higher solution rates in scientific literacy.

**Figure 10: Gender differences in science performance**

Source: PISA
Relative strengths

In the science domain too, relative strengths and weaknesses of students in Bhutan can be identified by applying a non-linear transformation to the solution rate on each item, then comparing the difference between this so-called “delta-score” and the average delta across all items with the corresponding difference in comparison countries.

Table 12: Relative strengths of Bhutan compared to the average of all PISA-D countries

<table>
<thead>
<tr>
<th>ITEM ID</th>
<th>ITEM NAME</th>
<th>BTN</th>
<th>PISA-D</th>
<th>DIFF</th>
<th>ITEM FORMAT</th>
<th>KNOWLEDGE</th>
<th>SYSTEM</th>
<th>COMPETENCY</th>
</tr>
</thead>
<tbody>
<tr>
<td>ps476q01s</td>
<td>Heart Surgery</td>
<td>11.24</td>
<td>14.00</td>
<td>-1.98</td>
<td>MCQ</td>
<td>Content</td>
<td>Living</td>
<td>Explain phenomena scientifically</td>
</tr>
<tr>
<td>ps269q03</td>
<td>Earth’s Temperature</td>
<td>15.75</td>
<td>18.34</td>
<td>-1.80</td>
<td>Open response</td>
<td>Content</td>
<td>Living</td>
<td>Explain phenomena scientifically</td>
</tr>
<tr>
<td>ps498q04</td>
<td>Experimental Digestion</td>
<td>14.21</td>
<td>16.73</td>
<td>-1.73</td>
<td>Open response</td>
<td>Procedural</td>
<td>Living</td>
<td>Interpret data and evidence scientifically</td>
</tr>
<tr>
<td>ps527q03as</td>
<td>Extinction of the Dinosaurs</td>
<td>11.33</td>
<td>13.77</td>
<td>-1.65</td>
<td>MCQ</td>
<td>Content</td>
<td>Earth and Space</td>
<td>Explain phenomena scientifically</td>
</tr>
<tr>
<td>ps256q01s</td>
<td>Spoons</td>
<td>8.35</td>
<td>10.77</td>
<td>-1.63</td>
<td>MCQ</td>
<td>Content</td>
<td>Physical</td>
<td>Explain phenomena scientifically</td>
</tr>
<tr>
<td>ps437q03s</td>
<td>Extinguishing Fires</td>
<td>13.35</td>
<td>15.67</td>
<td>-1.63</td>
<td>MCQ</td>
<td>Content</td>
<td>Physical</td>
<td>Explain phenomena scientifically</td>
</tr>
</tbody>
</table>

This method reveals 6 conspicuous items with slight differences in the delta scores of Bhutan compared to the average of all the PISA-D countries. The strengths of Bhutanese students can be seen primarily in items that require competency in explaining phenomena scientifically using content knowledge.

To classify the performance of Bhutan in more detail, the items identified in Table 12 are taken and we compare percent correct of Bhutan with the percent correct of these items in PISA-D and PISA reference countries.

Table 13: Relative strengths of Bhutan compared to the reference countries (PISA-D and PISA)

<table>
<thead>
<tr>
<th>ITEM ID</th>
<th>ITEM NAME</th>
<th>BTN</th>
<th>ECU</th>
<th>PRY</th>
<th>PISA-D AVG</th>
<th>SGP</th>
<th>SWE</th>
<th>DOM</th>
<th>OECD AVG</th>
</tr>
</thead>
<tbody>
<tr>
<td>ps476q01s</td>
<td>Heart Surgery</td>
<td>67.04</td>
<td>51.50</td>
<td>43.02</td>
<td>40.18</td>
<td>84.65</td>
<td>64.91</td>
<td>34.89</td>
<td>68.51</td>
</tr>
<tr>
<td>ps269q03</td>
<td>Earth’s Temperature</td>
<td>24.57</td>
<td>21.34</td>
<td>12.47</td>
<td>10.42</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>ps498q04</td>
<td>Experimental Digestion</td>
<td>28.10</td>
<td>31.51</td>
<td>15.77</td>
<td>18.55</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>ps527q03as</td>
<td>Extinction of the Dinosaurs</td>
<td>66.19</td>
<td>46.81</td>
<td>40.24</td>
<td>42.39</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>ps256q01s</td>
<td>Spoons</td>
<td>87.76</td>
<td>72.06</td>
<td>74.17</td>
<td>70.03</td>
<td>98.33</td>
<td>88.42</td>
<td>68.57</td>
<td>88.81</td>
</tr>
<tr>
<td>ps437q03s</td>
<td>Extinguishing Fires</td>
<td>46.48</td>
<td>40.99</td>
<td>26.89</td>
<td>25.70</td>
<td>57.54</td>
<td>49.32</td>
<td>23.92</td>
<td>49.40</td>
</tr>
</tbody>
</table>

Students in Bhutan performed significantly better compared to top PISA-D countries. On common PISA items, Bhutan’s solution rates are better than low performing PISA reference country and almost at par with OECD average.
Table 14: Relative weaknesses of Bhutan compared to the average of all PISA-D countries

<table>
<thead>
<tr>
<th>ITEM ID</th>
<th>ITEM NAME</th>
<th>BTN</th>
<th>PISA-D</th>
<th>DIFF</th>
<th>ITEM FORMAT</th>
<th>KNOWLEDGE</th>
<th>SYSTEM</th>
<th>COMPETENCY</th>
</tr>
</thead>
<tbody>
<tr>
<td>ps437q06</td>
<td>Extinguishing Fires</td>
<td>16.16</td>
<td>15.35</td>
<td>1.60</td>
<td>Open response</td>
<td>Content</td>
<td>Physical</td>
<td>Explain phenomena scientifically</td>
</tr>
<tr>
<td>ps252q01s</td>
<td>South Rainea</td>
<td>16.71</td>
<td>15.86</td>
<td>1.63</td>
<td>MCQ</td>
<td>Content</td>
<td>Earth and Space</td>
<td>Interpret data and evidence scientifically</td>
</tr>
<tr>
<td>ps408q05s</td>
<td>Wild Oat Grass</td>
<td>16.91</td>
<td>15.96</td>
<td>1.73</td>
<td>MCQ</td>
<td>Procedural</td>
<td>Living</td>
<td>Evaluate and design scientific enquiry</td>
</tr>
<tr>
<td>ps415q07s</td>
<td>Solar Panels</td>
<td>14.23</td>
<td>13.13</td>
<td>1.89</td>
<td>MCQ</td>
<td>Epistemic</td>
<td>Earth and Space</td>
<td>Evaluate design scientific enquiry</td>
</tr>
<tr>
<td>ps7221q1as</td>
<td>Clean Drinking Water</td>
<td>9.29</td>
<td>8.04</td>
<td>2.03</td>
<td>MCQ</td>
<td>Content</td>
<td>Living</td>
<td>Explain phenomena scientifically</td>
</tr>
<tr>
<td>ps438q03</td>
<td>Green Parks</td>
<td>19.40</td>
<td>18.07</td>
<td>2.11</td>
<td>Open response</td>
<td>Epistemic</td>
<td>Physical</td>
<td>Evaluate and design scientific enquiry</td>
</tr>
<tr>
<td>ps413q05s</td>
<td>Plastic Age</td>
<td>14.65</td>
<td>13.25</td>
<td>2.19</td>
<td>MCQ</td>
<td>Content</td>
<td>Physical</td>
<td>Interpret data and evidence scientifically</td>
</tr>
</tbody>
</table>

Table 15: Relative weaknesses of Bhutan compared to the reference countries (PISA-D and PISA)

<table>
<thead>
<tr>
<th>ITEM ID</th>
<th>ITEM NAME</th>
<th>BTN</th>
<th>ECU</th>
<th>PRY</th>
<th>PISA-D AVG</th>
<th>SGP</th>
<th>SWE</th>
<th>DOM</th>
<th>OECD AVG</th>
</tr>
</thead>
<tbody>
<tr>
<td>ps437q06</td>
<td>Extinguishing Fires</td>
<td>21.50</td>
<td>39.35</td>
<td>34.66</td>
<td>28.83</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ps252q01s</td>
<td>South Rainea</td>
<td>17.71</td>
<td>31.81</td>
<td>33.22</td>
<td>24.15</td>
<td>53.91</td>
<td>58.02</td>
<td>18.48</td>
<td>52.94</td>
</tr>
<tr>
<td>ps408q05s</td>
<td>Wild Oat Grass</td>
<td>16.44</td>
<td>30.58</td>
<td>24.52</td>
<td>23.41</td>
<td>44.40</td>
<td>41.49</td>
<td>12.92</td>
<td>39.17</td>
</tr>
<tr>
<td>ps415q07s</td>
<td>Solar Panels</td>
<td>37.97</td>
<td>59.83</td>
<td>54.59</td>
<td>48.79</td>
<td>67.62</td>
<td>69.66</td>
<td>36.42</td>
<td>73.61</td>
</tr>
<tr>
<td>ps7221q1as</td>
<td>Clean Drinking Water</td>
<td>82.31</td>
<td>97.97</td>
<td>93.28</td>
<td>84.51</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ps438q03</td>
<td>Green Parks</td>
<td>5.48</td>
<td>13.36</td>
<td>9.66</td>
<td>11.37</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ps413q05s</td>
<td>Plastic Age</td>
<td>34.01</td>
<td>58.13</td>
<td>56.01</td>
<td>47.65</td>
<td>66.91</td>
<td>60.53</td>
<td>37.26</td>
<td>67.58</td>
</tr>
</tbody>
</table>

The contrast between strengths and weaknesses in Table 12 and 14 suggests that students in Bhutan may find items testing competency evaluate and design scientific enquiry and epistemic knowledge particularly difficult.

Although Bhutan performed better than most PISA-D countries, there is a huge performance gap between Bhutan and PISA countries. Bhutan’s performance is about 23 to 35 points (percentage points) below OECD averages.

Table 16: Average deltas for science item classification of PISA-D countries

<table>
<thead>
<tr>
<th>ITEM GROUP</th>
<th>BTN</th>
<th>ECU</th>
<th>GTM</th>
<th>HND</th>
<th>KHM</th>
<th>PRY</th>
<th>SEN</th>
<th>ZMB</th>
<th>PISA-D AVG</th>
</tr>
</thead>
</table>
Analysis of illustrating items

In this section, we will illustrate some of the patterns observed in Table 12 and 14 with a typical PISA release item. Each item in the PISA test has its own properties and its specific cognitive requirements.

A. Strength patterns

To illustrate the pattern, we have chosen item “Mary Montagu” and “Bee Colony Collapse Disorder”. These items’ characteristics are similar to the weakness patterns observed for Bhutan.

**Item 1** Mary Montagu

Read the following newspaper article and answer the questions that follow.

**The History of Vaccination**

Mary Montagu was a beautiful woman. She survived an attack of smallpox in 1715 but she was left covered with scars. While living in Turkey in 1717, she observed a method called inoculation that was commonly used there. This treatment involved scratching a weak type of smallpox virus into the skin of healthy young people who then became sick, but in most cases only with a mild form of disease.

Mary Montagu was so convinced of the safety of these inoculations that she allowed her son and daughter to be inoculated.

In 1796, Edward Jenner used inoculations of a related disease, cowpox, to produce antibodies against smallpox. Compared with the inoculation of smallpox, this treatment had less side effects and the treated person could not infect others. The treatment became known as vaccination.

**Question**

What kinds of diseases can people be vaccinated against?

A. Inherited diseases like haemophilia.

B. Diseases that are caused by viruses, like polio.

C. Diseases from the malfunctioning of the body, like diabetes.
D. Any sort of disease that has no cure.

Ans: ___________________________________________

Source: OECD 2014 What Students Know and Can Do: Student Performance in Mathematics, Reading and Science – Volume I

This task requires student to recall a specific piece of knowledge that vaccination helps prevent diseases, the cause for which is external to normal body components. This fact is then applied in the selection of the correct explanation and the rejection of other explanations. The term “virus” appears in the stimulus text and provides a hint for students. This has lowered the difficulty of the question. Recalling an appropriate, tangible scientific fact and its application in a relatively simple context locates the question at Level 2.

Item 2  Bee Colony Collapse Disorder

This item requires students to provide a hypothesis for the collapses among the control bee colonies. A correct response indicates either that there must be another natural cause of colony collapse for the studied colonies or that the hives in the control group were not properly protected from exposure.

<table>
<thead>
<tr>
<th>Competency</th>
<th>Explain Phenomena Scientifically</th>
</tr>
</thead>
<tbody>
<tr>
<td>Knowledge – System</td>
<td>Content – Living</td>
</tr>
</tbody>
</table>
B. Weakness patterns

To illustrate the pattern, we have chosen item “Slope-Face Investigation” and “Running in Hot Weather”. These items’ characteristics are similar to the weakness patterns observed for Bhutan.

**Item 1  Slope Face Investigation**

![Slope Face Investigation](http://www.oecd.org/pisa/test/PISA%202015%20MS%20-%20Released%20Item%20Descriptions%20Final_English.pdf)

The item requires students to apply epistemic knowledge to explain the design of the investigation presented in this unit. This Level 3 question allows students to demonstrate their understanding of the underlying rationale for the procedure of taking two independent measures of the phenomena being investigated. Knowledge of this rationale is the aspect of this question that assesses epistemic knowledge.

**Item 2  Running in Hot Weather**
In this item, students are asked to use the simulation to identify the highest temperature at which a person can run without getting heat stroke when the humidity is 40%. The correct response is 35°C and students must select the following two rows of data to support their response: 35°C air temperature - 40% humidity and 40°C air temperature - 40% humidity. They must further explain how the selected rows of data support their answer by indicating that at 40% humidity moving the air temperature up from 35°C to 40°C causes heat stroke.

<table>
<thead>
<tr>
<th>Competency</th>
<th>Evaluate and Design Scientific Enquiry</th>
</tr>
</thead>
<tbody>
<tr>
<td>Knowledge - System</td>
<td>Procedural</td>
</tr>
<tr>
<td>Context</td>
<td>Personal – Health and Disease</td>
</tr>
<tr>
<td>Cognitive Demand</td>
<td>Medium (Level 4)</td>
</tr>
<tr>
<td>Item Format</td>
<td>Open Response – Human Coded</td>
</tr>
</tbody>
</table>

In this item, one variable is defined. With a set air humidity of 40%, students must run at least two trials in order to determine the highest temperature at which a person can run without getting heat stroke. They must draw on procedural knowledge to explain how the data they have collected supports their answer by indicating that at 40% humidity, an air temperature higher than 35°C results in heat stroke.
Looking forward: Policy options for Bhutan
The Royal Government of Bhutan accords highest priority to education sector as the country’s quality of health, prosperity, happiness and progression hinges on the quality of its education. Towards this effect, the Ministry of Education has taken several reform initiatives to ensure that there are improvements in access, equity and system efficiencies to improve the quality of education in the country. Over the past decades, Bhutan has made rapid progress in the education sector, and the country has immensely benefitted from a generation of nation builders the system produced, however, the systems still face challenges in delivering quality education.

Ministry of Education has undertaken an ambitious task to participate in the international benchmarking systems, such as PISA-D and PISA, to report on the state of education in the country. With participation in PISA-D, ministry seeks to set a benchmark profile of the knowledge, skills and competencies of the students in Bhutan and to collect evidences on the readiness of Bhutanese education system for the participation in the PISA 2021.

Summary of findings

Overall performance

Bhutan’s performance in PISA-D is remarkable similar to low and middle-income countries. Although no PISA scores were generated for the Bhutan due to small sample size, a reliable estimate based on the percent correct scores shows students’ performance significantly below OECD average. And when compared to other PISA for development countries, Bhutanese students’ performance lies in-between the two highest performing PISA-D countries.

Summary of relative strengths in the three domains

In reading literacy, Bhutanese students had signif cantly higher solution rates in items related to access and retrieve aspect compared to PISA-D reference countries. In such items, students are required to locate individual pieces of information, such as the details required by an employer from a job advertisement, to find a telephone number with several prefix codes, to find a particular fact to support or disprove a claim someone has made. Majority of conspicuous items were based on personal situations and of open response type.

In mathematical literacy, Bhutanese students had higher solutions rates in items related to “Employing Mathematical Concepts, Facts and Procedures” and “Interpreting, Applying and Evaluating Mathematical Outcomes” compared to PISA-D reference countries. These items were from mathematical process involving “Change and Relationships” and “Uncertainty and Data” mostly based on societal situations.
In scientific literacy, Bhutanese students performed significantly well in items testing “Explain phenomena scientifically” competency compared to PISA-D countries. Majority of these items tested content knowledge related to living systems. These findings in the three domains with the conspicuous items are consistent with the overall summary statistics.

Summary of relative weaknesses in the three domains

In reading literacy, about 61% of the items had a delta value of higher than 13, that is, students found these items difficult. Two significantly difficult aspects were “Reflect and evaluate” and most of the items were related to “Public” situations with delta values of 15.29 and 14.90 respectively. These items require students to use their own experience or knowledge to compare, contrast or hypothesise and make a judgment drawing on standards beyond the text. Much of these text are associated with out-of-school settings for students, such as rules for clubs and records of games, which often take place unofficially at school. Such items are found more difficult as it requires students to draw on narrow, specialised knowledge rather than broad and common knowledge.

In mathematical literacy, about 70% of the items had a delta value higher than 13, that is, about seven out of ten items. Items related to the mathematical process, formulating situations mathematically were the most difficult ones with an average delta score of 16.94. Although, students had higher solution rates in items related to mathematical content knowledge “Change and relationship” but overall, they found the items difficult. Students also found the items related to mathematical context “Scientific” comparatively more difficult than the other three contexts. In change and relationship, the average delta was 15.51 and in scientific context the delta was 15.21. These tasks involve translating mathematical solutions or reasoning back into the context of a problem and determining whether the results are reasonable and make sense in the context problem. Much of these items are related to the application of mathematics to the natural world and issues and topics related to science and technology, such as (but not limited to), weather or climate, ecology, medicine, space science, genetics, measurement and the world of mathematics itself. Tasks also require modeling the change and the relationships with appropriate functions and equations, as well as creating, interpreting, and translating among symbolic and graphical representations of relationships.

In scientific literacy, about 58% of the items had a delta value higher than or equal to 13. Items related to epistemic knowledge were more difficult than the other two types of knowledge. And the items dealing with the competency, “Evaluate and design scientific enquiry” were comparatively difficult than the other two with an average delta score of 14.28. These tasks requires students to identify whether the conclusions are justified by the data, or what piece of evidence best supports the hypothesis advanced in an item and explain why. Some of these tasks required students to evaluate reports of scientific findings and investigations critically. These tasks relied on students’ ability to distinguish scientific questions from other forms of enquiry or recognise questions that could be investigated scientifically in a given context.
Implications of the strengths and weaknesses for Bhutan

The strengths and weaknesses for Bhutan are very similar to the other PISA-D and low-income countries. Students in general have higher success rates in items requiring lower cognitive skills, however, there is a significant gap in performance in more demanding tasks. This also reflects to some extent the fact that the kind of knowledge and understanding demonstrated by Bhutan’s students tend to be relatively broad, in comparison to other PISA-D countries, but shallow.

In the following section, we will highlight PISA-D and general findings which have implications on students’ performance. There is also need to train teachers to develop deep understanding and encourage students to actively engage with the subject and with problem solving. Although, basic skills mentioned are important it is very clear that, focus on higher order cognitive skills is vital in preparing students for the future.

- Ministry of Education / Schools / Royal Education Council / Bhutan Council for School Examinations and Assessment should strengthen and enhance competency based activities and assessment in curriculum; teachers’ knowledge and skills in competency based teaching and learning and teachers’ knowledge and skills in developing competency based items.
- Bhutan Council for School Examinations and Assessment should strengthen competency based items in high stake examinations.
- Teacher training colleges should incorporate competency based teaching and learning modules in pre-service training programme;
- Schools should enhance and strengthen school-based assessment;
- Ministry of Education should strengthen the monitoring of quality instructions through the existing school self-assessment tool and school improvement plan.
- Schools / Royal Education Council / Bhutan Council for School Examinations and Assessment should prioritize on the depth rather than the breadth of learning to avoid superficial learning, incomplete understanding of core concepts and limited ability to transfer and apply knowledge to unfamiliar contexts.
- Schools should provide ICT facilities, better Internet connectivity and promote integration of ICT in teaching-learning process.
- School should be empowered in decision-making, budgeting and hiring and dismissing of teachers.
- Schools should strengthen and enhance bespoke school-based professional development.
- Schools should emphasize on teacher appraisal mechanism on improving teachers’ instructional and classroom management practices.
- Teachers should identify the students at risk of failing in academics and provide additional support to avoid grade repetition.
- Parents and teachers should break the gender stereotypes about science / mathematics / reading related activities and occupations to allow children achieve their full potentials.
PISA findings which have implications on students’ performance in the three domains are as follows:

Mathematical Literacy
- Mathematics teachers should use diverse teaching strategies considering both the content and the learners with different abilities, motivation and interests. The focus should be on student-centered teaching rather than teacher-centered teaching.
- Apart from teaching fundamental elements of the mathematics curriculum, mathematics teachers should expose students to a wide range of problems and contexts to better adapt to the changing needs of the world.
- Mathematics teachers should use metacognition activation strategy to encourage students to think about their own learning which can help them monitor progress and reveal their own learning difficulties.
- Mathematics teachers should be provided with high quality professional development in subject knowledge and understanding, pedagogical competencies and processes for improving their confidence, abilities and skills to deliver quality mathematical instructions.

Scientific Literacy
- Science teachers should emphasize more on deep conceptual and epistemic knowledge of science curricula and instructions.
- Science teachers should make teaching more effective by having a mix of inquiry-based and teacher-directed instructions. The special focus should be given on explaining scientific ideas, demonstrating ideas and adapting the lesson to the students’ needs and knowledge. Science teachers should provide individual help when a student has difficulties understanding a topic or questions.
- Schools should ensure adequate laboratory materials coupled with well structured laboratory activities to provide hands-on activities and experience to students to promote critical thinking.
- Schools should offer science competitions, clubs and other extracurricular activities to help students understand scientific concepts, raise interest in science and even nurture future scientists.

Reading Literacy
- Schools should institute reading culture; strengthen teacher model reading; offer literary activities and provide varieties of reading materials to enhance student’s reading and comprehension skills which are important life skills rather than an exercise in English subject.
- Parents should motivate their children at home by reading books, discussing on political issues, social issues, films, music and other cultural events to allow children to develop informed opinions and help improve their critical thinking.
- Teachers should stay abreast of the curricular learning outcomes and evolving pedagogies in reading to identify and incorporate higher order thinking skills in teaching-learning process.
Next steps in Bhutan’s participation in PISA 2021

Given the nature of Bhutan’s participation in PISA for Development, in-depth understanding of the nature of strengths and weakness cannot be ascertained on a significant level to be able to recommend focused measures and strategies for improvement, however with Bhutan’s participation in PISA 2021, full insight that will help the students in obtaining the knowledge and skills needed to succeed in the tomorrow’s world can be realized.
REFERENCES


» Australian Council for Educational Research Ltd (2017) TIMSS 2015 Reporting Australia’s results, 19 Prospect Hill Road, Camberwell, Victoria, 3124, Australia

» BCSEA (2015), BCSEA Journal for Educational Assessment, Babesa Thimphu Bhutan


» Ministerial Advisory Group on Literacy and Numeracy, NSW Department of Education and Communities February 2013


» OECD (2016), PISA 2015 Results (Volume I): Excellence and Equity in Education, PISA,

» EDUCATION AND SKILLS Measuring Student Knowledge and Skills. THE PISA 2000 ASSESSMENT OF READING, MATHEMATICAL AND SCIENTIFIC LITERACY

» OECD (2017) PISA for Development Assessment and Analytical Framework READING, MATHEMATICS AND SCIENCE
