



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

Findings from Bhutan's Experience in PISA for Development (PISA-D)

Published by National Project Centre, Bhutan Council for School Examinations and Assessment, Thimphu, Bhutan

© 2019, Bhutan Council for School Examinations and Assessment

Website: www.bcsea.bt

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior permission of publisher.

This report was developed by Bhutan jointly with the Organisation for Economic Co-operation and Development (OECD), which assisted in the design of the report and provided input, guidance and assistance in its development. The report is published under the responsibility of Bhutan and does not necessarily represent the official views of the OECD or its member countries.

ISBN 978-99936-718-8-6

ACKNOWLEDGEMENT

Bhutan 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, principals and parents.

We would like to acknowledge officials from the 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.

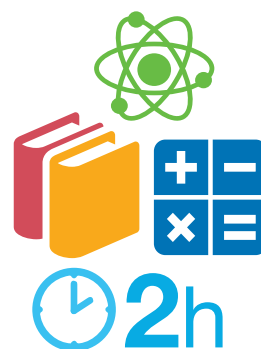
We are thankful to the OECD, Education Testing Service and Westat for the technical support and guidance provided at various stages of the project and also grateful to OECD for designing the template and finalizing the Bhutan PISA-D National Report.



In November 2017, **2,457 students** representing **10,725 15-year-olds** ...



from 53 schools representing **205 secondary schools** ...



...took a **2-hour test** in Reading, Mathematics and Science.

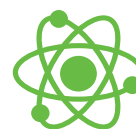
In Bhutan, average solution rate in



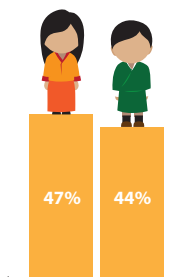
Reading literacy is 45.3%.



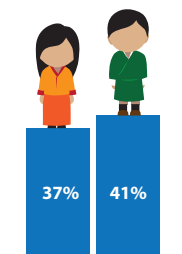
Mathematical literacy is 38.8%.



Scientific literacy is 45.1%.



Girls outperform boys in reading by 3 percentage points.



Boys outperform girls in mathematics by 4 percentage points.



Boys outperform girls in science by 1 percentage point.

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 and of the best education systems in Asia.

CONTENTS

CHAPTER 1	PISA AND PISA FOR DEVELOPMENT	1
	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	9
	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 PARTICIPATION IN THE FUTURE CYCLE OF PISA	45
	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 FUTURE CYCLES OF PISA	49

FOREWORD

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 basic 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 Bhutan Council for School Examinations and Assessment and Ministry of Education 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 our existing 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: 45.34% in Reading literacy, 38.84% in Mathematics literacy and 45.10% 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 intends 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 like to commend our Bhutan PISA-D National Report authors; Arjun Kumar Gurung (Deputy National Project Manager and Domain Expert Mathematical literacy), Karma Jigme Lepcha (Lead Analyst), Kinley Dema (Domain Expert Reading literacy), Sonam Lhamo (Domain Expert Scientific literacy) and Pem Tshering (PISA-D Focal Officer, Ministry of Education).



(Jai Bir Rai)
Minister
Ministry of Education
Bhutan

PREFACE

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.

ACRONYMS

BCSEA	Bhutan Council for School Examinations and Assessment
BTN	Bhutan
DOM	Dominican Republic
ECU	Ecuador
EMIS	Education Management and Information Service
ETS	Education Testing Service
GNH	Gross National Happiness
GTM	Guatemala
HND	Honduras
ICT	Information and Communications Technology
KHM	Cambodia
MoE	Ministry of Education
NEA	National Education Assessment
NPC	National Project Centre
OECD	Organisation for Economic Co-operation and Development
PAN	Panama
PCE	Paro College of Education
PISA	Programme for International Student Assessment
PISA-D	PISA for Development
PRY	Paraguay
REC	Royal Education Council
SDG	Education Sustainable Development Goal
SEN	Senegal
SGP	Singapore
SWE	Sweden
ZMB	Zambia



Bhutan in PISA-D

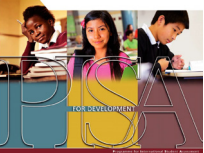
In Bhutan during the month of November 2017 more than 2,400 15-year-old students in class VII (7th grade) or above from 53 randomly selected schools across the country took a two-hour test in reading, mathematics and science. These tests were not directly linked to Bhutan's school curriculum – rather, 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-D 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-D 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 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-D 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 (Education Management and Information System, 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.

As the sample for Bhutan was not designed to provide representative results for the country, the results of the PISA-D participation of Bhutan refer, 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

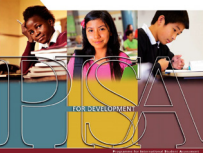
Over the past two decades, PISA has steadily increased the number of participating countries, from 44 in 2000 to 82 in 2018. As the number of countries joining PISA increases, PISA evolves to successfully cater for a larger and more diverse group of participants. Bhutan decided to join the PISA for Development (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 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 is 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 in PISA-D contain items that are 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 low- and middle-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

Bhutan participated in PISA-D to gain experience from this initiative to help 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

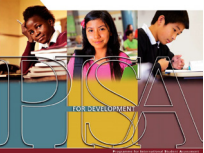
This national report and other related communication products present Bhutan's results in context to the countries that participated in PISA 2015 and PISA-D and include relevant analyses and information based on the policy priorities of Bhutan. This report constitutes a summary of key results and analysis 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 Bhutan's students 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's students 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-D test was statistically confirmed, the average solution rates across items were examined to describe the performance of the students who took the PISA-D 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 Bhutan's students 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 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-D 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 Bhutan's 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's 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 three domains. The solution rates for Bhutan are in between those of the two highest performing PISA-D countries (Ecuador and Paraguay) which are therefore selected as reference countries.

Table 1: Percent-correct averages for PISA-D countries

DOMAIN	BTN	ECU	GTM	HND	KHM	PRY	SEN	ZMB	PISA-D AVG
Reading literacy	45.34	52.66	44.50	44.74	34.62	47.03	37.37	27.88	41.26
Mathematical literacy	38.84	41.69	31.33	33.21	30.29	31.44	27.52	19.08	30.65
Scientific literacy	45.10	46.98	39.82	40.71	34.85	41.16	33.25	31.19	38.28

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's performance in PISA-D in more detail in each of the three domains tested: reading, mathematics and science.

Reading literacy assessment in PISA-D

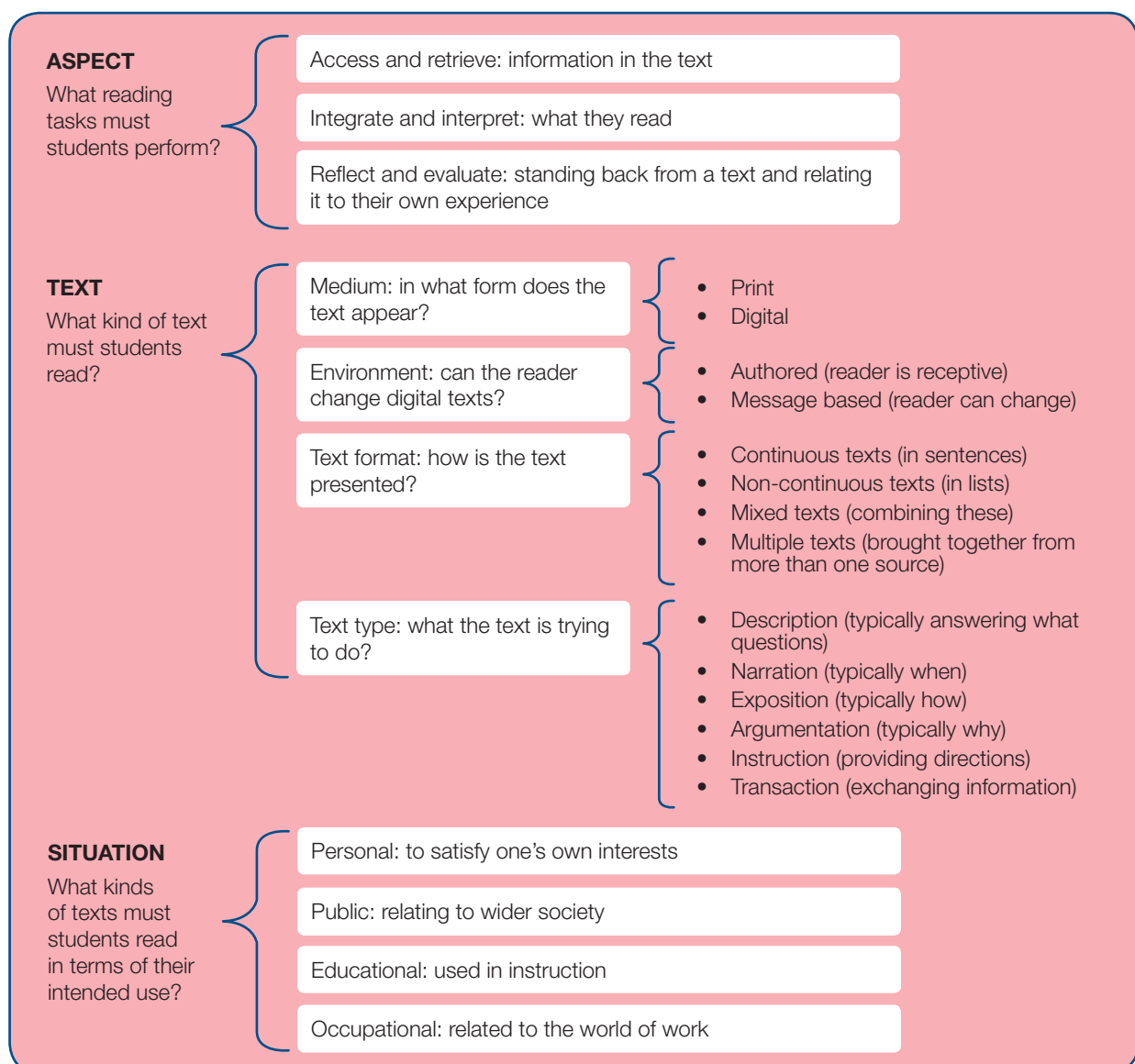
Reading literacy is defined as an individual's capacity to understand, use, reflect on and engage with written texts, in order to achieve one's goals, to develop one's knowledge and potential, and to participate in society.

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)

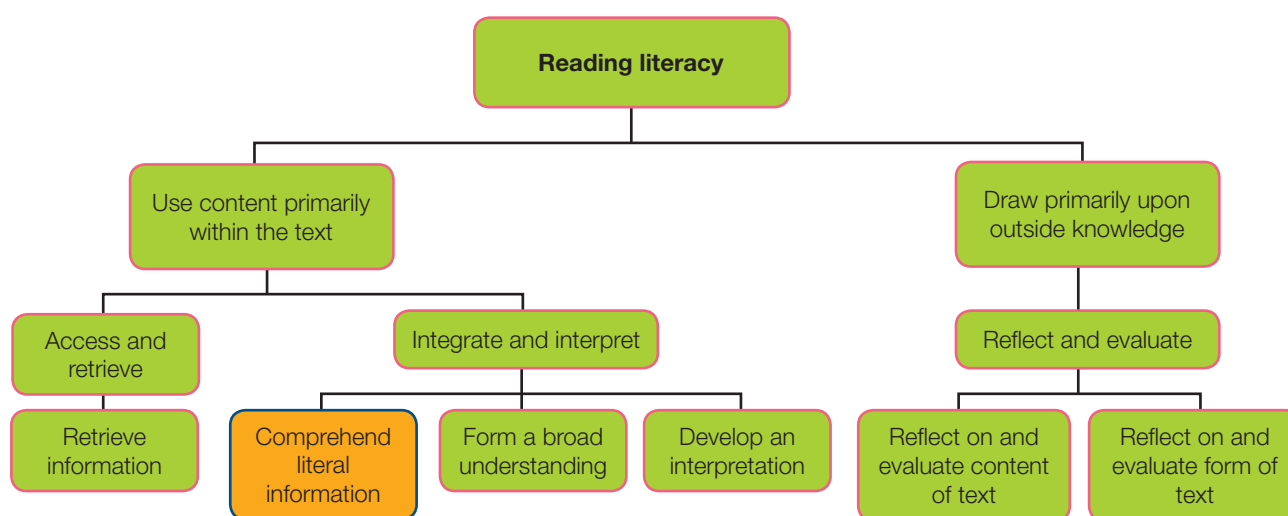




For PISA-D, an additional process titled “literal comprehension” has been added. Literal comprehension requires students to comprehend explicitly stated information that may be found in individual words, sentences or passages. In addition, the concept of “retrieving information” is broadened to range 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.

Figure 2: Processes targeted in Reading literacy test development for PISA and PISA-D



Solution rates in Reading in Bhutan

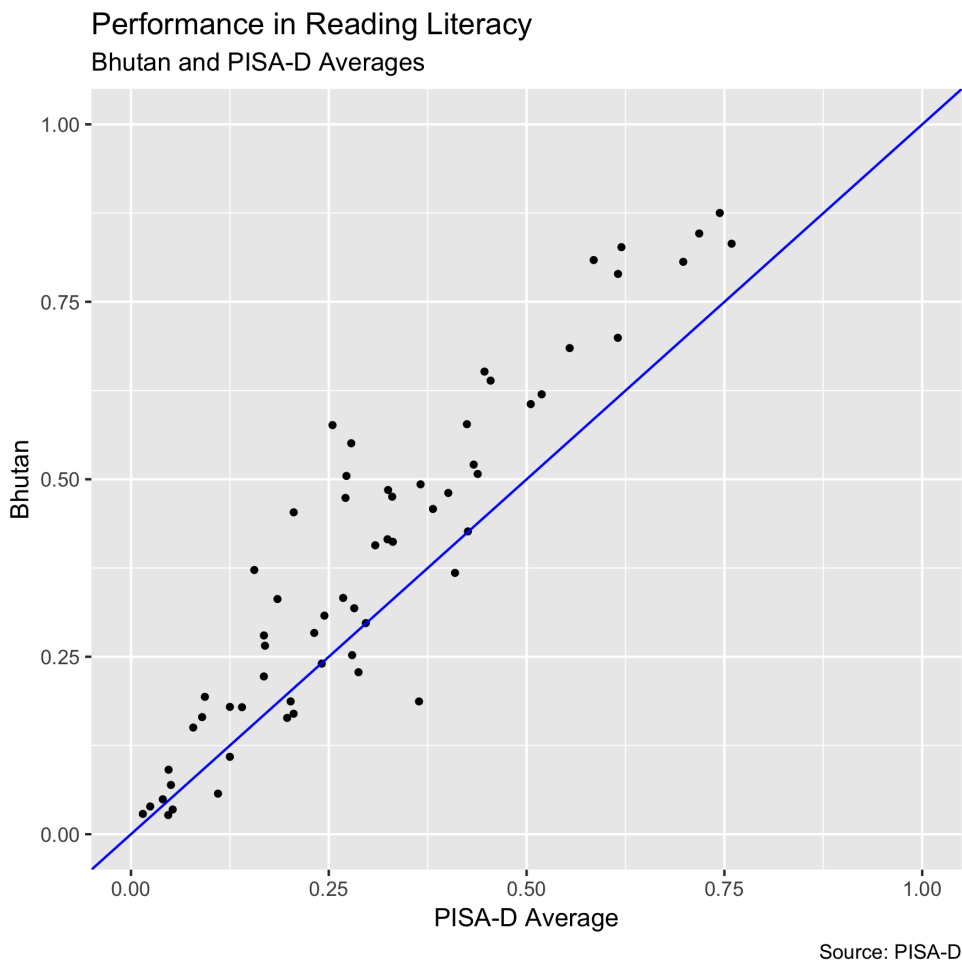
Bhutan’s students achieved an average solution rate of 45.34 percent in the PISA-D 2017 Reading literacy assessment, which was higher than the PISA-D average solution rate of 41.26 percent.

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.

Overall performance and gender differences

Figure 3 visually compares the solution rates of students who took the PISA-D Reading 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 achieved the same solution rate as students in comparison countries on average. The dots below the blue line are the items in which Bhutan’s students had lower solution rates than PISA-D average.

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



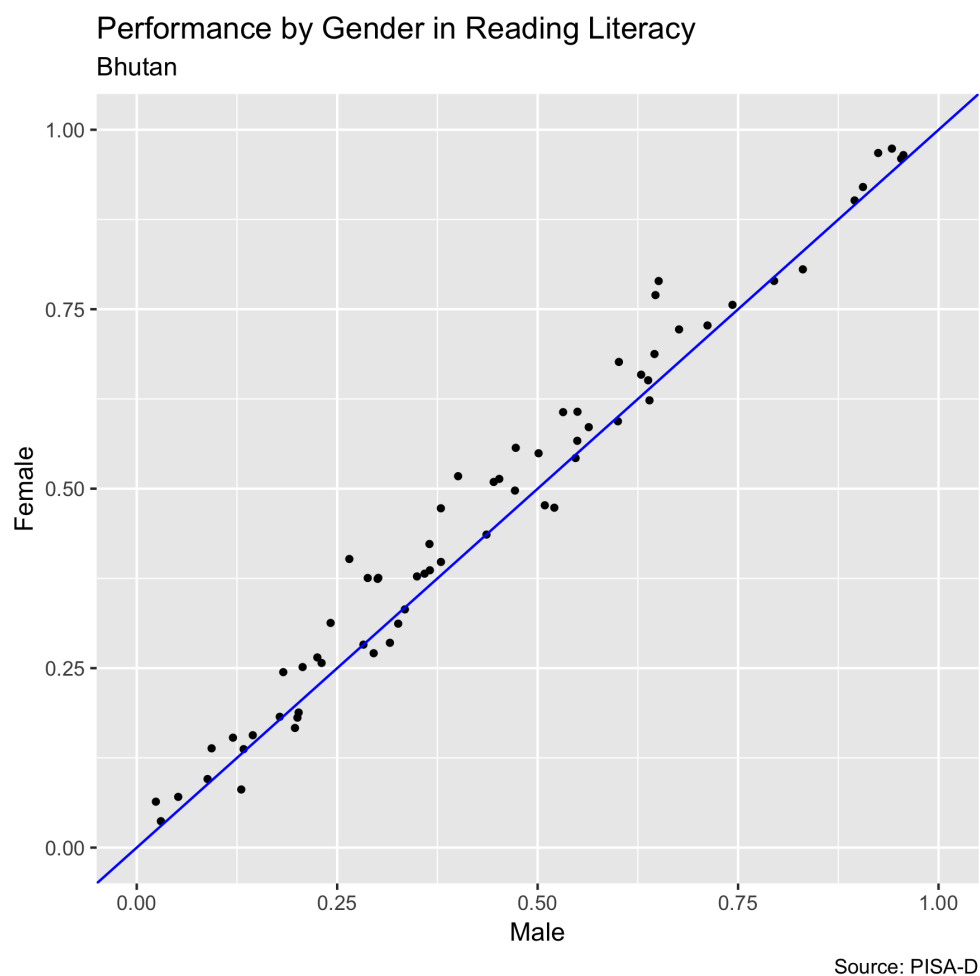
Percent-correct scatter plot, Bhutan vs PISA-D average (Reading literacy)

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 Bhutan's 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 blue diagonal line represents a benchmark situation in which both boys and girls in Bhutan achieved the same solution rate. 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 literacy.



Figure 4: Gender differences in reading performance

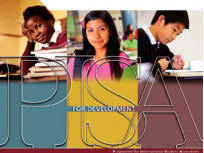


Percent-correct scatter plot, Boys vs Girls

Relative strengths

Relative strengths of Bhutan's students are items on which the solution rates are 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

ITEM ID	ITEM NAME	DELTA			ITEM FORMAT	SITUATION	ASPECT
		BTN	PISA-D	DIFF			
pr460q01	Gulf of Mexico	10.65	13.80	-2.62	Open response	Educational	Access and retrieve
pr442q02	Galileo	9.39	12.02	-2.09	Open response	Personal	Access and retrieve
pr455q02	Chocolate and Health	15.76	17.98	-1.68	Open response	Personal	Access and retrieve
pr6018q3a	Aski Gym Youth Programme	15.94	18.08	-1.60	Open response	Personal	Integrate and interpret
pr456q01s	Biscuits	5.95	8.09	-1.60	Multiple choice	Personal	Access and retrieve

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.

These items constitute relative strengths for Bhutan which 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-D and OECD averages.

Table 3: Relative strengths of Bhutan compared to the reference countries (PISA-D and PISA)

ITEM ID	ITEM NAME	BTN	ECU	PRY	PISA-D AVG	SGP	SWE	DOM	OECD AVG
pr460q01	Gulf of Mexico	72.16	53.74	53.17	42.23	NA	NA	NA	NA
pr442q02	Galileo	81.67	77.73	72.97	59.19	NA	NA	NA	NA
pr455q02	Chocolate and Health	24.47	23.84	17.61	12.99	NA	NA	NA	NA
pr6018q3a	Aski Gym Youth Programme	23.08	15.45	16.08	11.25	NA	NA	NA	NA
pr456q01s	Biscuits	96.10	96.02	92.61	87.48	97.57	92.63	80.82	95.35

Bhutan’s students 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 6 shows that, of the three aspects, Bhutan’s 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

ITEM ID	ITEM NAME	DELTA			ITEM FORMAT	SITUATION	ASPECT
		BTN	PISA-D	DIFF			
pr432q01	About a book	14.32	13.24	1.62	Open response	Personal	Integrate and interpret
pr442q05	Galileo	19.73	18.56	1.70	Open response	Personal	Reflect and evaluate
pr101q03s	Rhino	16.15	14.60	2.09	MCQ	Public	Reflect and evaluate
pr101q02s	Rhino	14.39	11.47	3.45	MCQ	Public	Integrate and interpret
pr101q01s	Rhino	18.03	14.90	3.67	MCQ	Public	Integrate and interpret

Table 4 displays 5 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 Bhutan's 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 4 are taken in Table 5 to 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 (PISA-D and PISA)

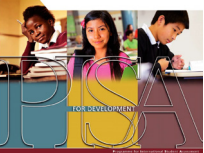
ITEM ID	ITEM NAME	BTN	ECU	PRY	PISA-D AVG	SGP	SWE	DOM	OECD AVG
pr432q01	About a book	37.05	56.01	51.13	47.91	NA	NA	NA	NA
pr101q03s	Rhino	21.56	51.75	43.44	35.18	65.05	70.48	NA	64.58
pr442q05	Galileo	4.63	13.88	18.07	10.35	NA	NA	NA	NA
pr101q02s	Rhino	36.45	77.71	75.59	64.26	80.90	82.86	NA	84.59
pr101q01s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

The strength of Bhutan's students mostly manifest on easier items, for example, items in strengths had delta values from 5.95 to 15.94. And weaknesses 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

ITEM GROUP	BTN	ECU	GTM	HND	KHM	PRY	SEN	ZMB	PISA-D AVG
Access and retrieve	11.21	10.51	11.49	11.41	12.59	11.32	12.80	14.07	12.03
Integrate and interpret	14.39	13.46	14.37	14.27	15.68	13.88	14.94	16.21	14.69
Reflect and evaluate	15.29	14.10	15.38	15.60	16.87	15.22	16.12	18.07	15.91
Educational	13.02	12.32	13.24	13.37	14.64	13.12	14.46	15.46	13.80
Occupational	12.01	11.21	12.28	11.73	11.81	11.80	12.65	13.47	12.14



ITEM GROUP	BTN	ECU	GTM	HND	KHM	PRY	SEN	ZMB	PISA-D AVG
Personal	13.06	12.21	13.32	13.16	14.79	12.91	14.04	15.93	13.77
Public	14.90	13.68	14.67	14.72	15.97	14.29	15.38	16.79	15.07
Selected-response	13.84	13.02	13.77	13.68	14.96	13.57	14.52	15.30	14.12
Open and close constructed-response	13.29	12.30	13.53	13.54	14.88	13.12	14.46	16.37	14.03

Analysis of illustrating items

In this section, we will illustrate some of the patterns observed in Table 2 and 4 with a typical PISA released items. 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 strength 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 students are required to access and retrieve a piece of explicitly stated information. In order to answer the question, students have 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: students 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).



Item 2 The Galapagos Islands

PISA 2018

The Galapagos Islands
Question 3 / 7

Refer to the different webpages of the website on the right.
Click on a choice to answer the question.

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 – A NATURAL TREASURE


About **Animals** **Conservation** **Volunteer**

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.

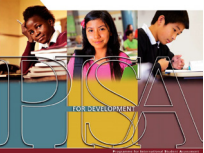

The Galapagos Islands


South America

This item does not explicitly identify the relevant text for students. 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 students 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.

Cognitive Process	Access and retrieve information within a text
Response Format	Simple multiple choice



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

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

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.

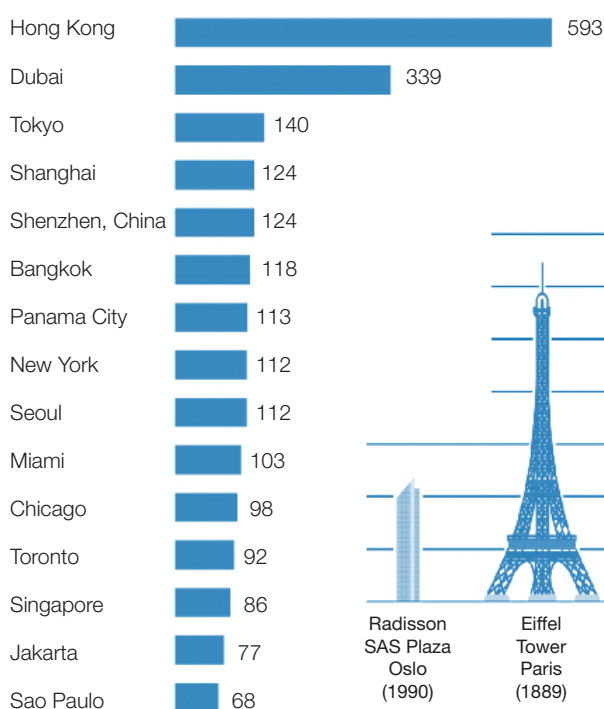
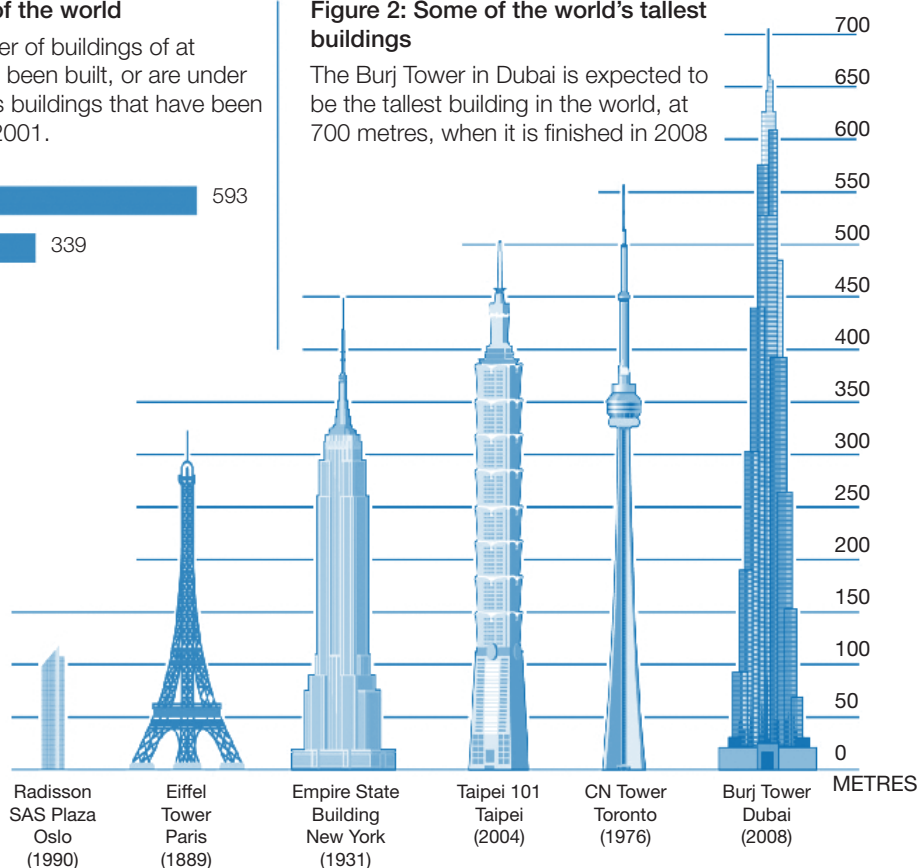


Figure 2: Some of the world's tallest buildings

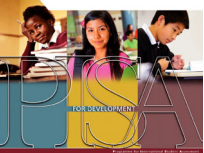
The Burj Tower in Dubai is expected to be the tallest building in the world, at 700 metres, when it is finished in 2008.



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

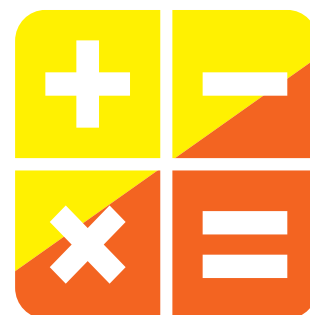
Situation	Educational
Medium	Print
Text format	Non-continuous
Text type	Exposition
Aspect	Reflect and evaluate: Reflect on and evaluate the content of a text
Question intent	Recognise the influence of reader's perspective on the way a text is constructed
Item format	Open constructed response



This item was difficult. Only about one-quarter of students answered correctly. Students are required 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

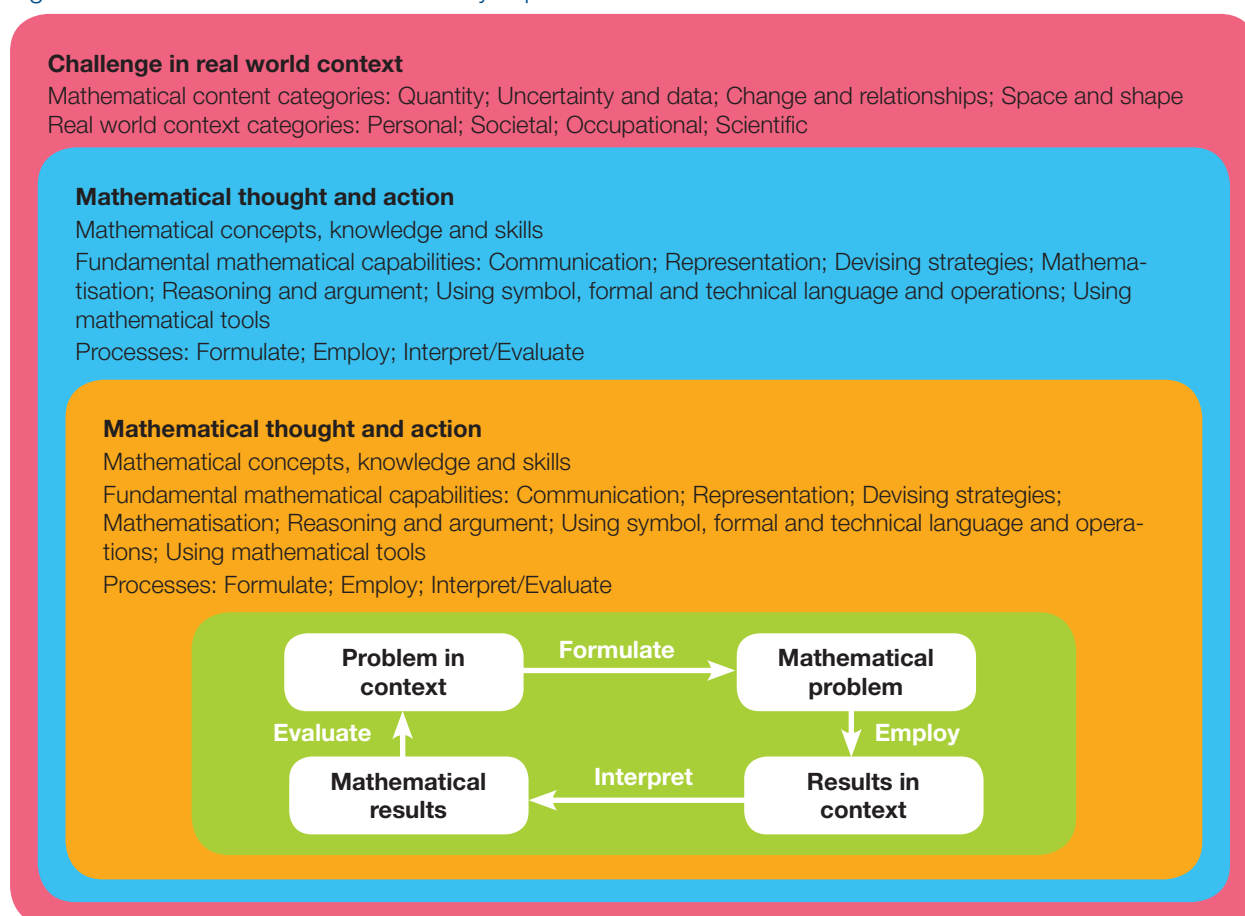
Mathematical literacy is defined as an individual's capacity to formulate, employ, and interpret mathematics in a variety of contexts. It includes reasoning mathematically and using mathematical concepts, procedures, facts and tools to describe, explain and predict phenomena. It assists individuals to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



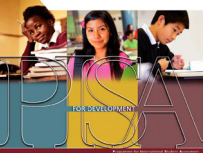
In mathematics, the baseline level of skills is defined as the level at which students can not only carry out routine procedures, such as an arithmetic operation, in situations where all the instructions are given to them, but can also interpret and recognise how a (simple) situation (e.g. comparing 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



Source: OECD, 2013(a), PISA 2012 Assessment and Analytical Framework: Mathematics, Reading, Science, Problem Solving and Financial Literacy, OECD Publishing.



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

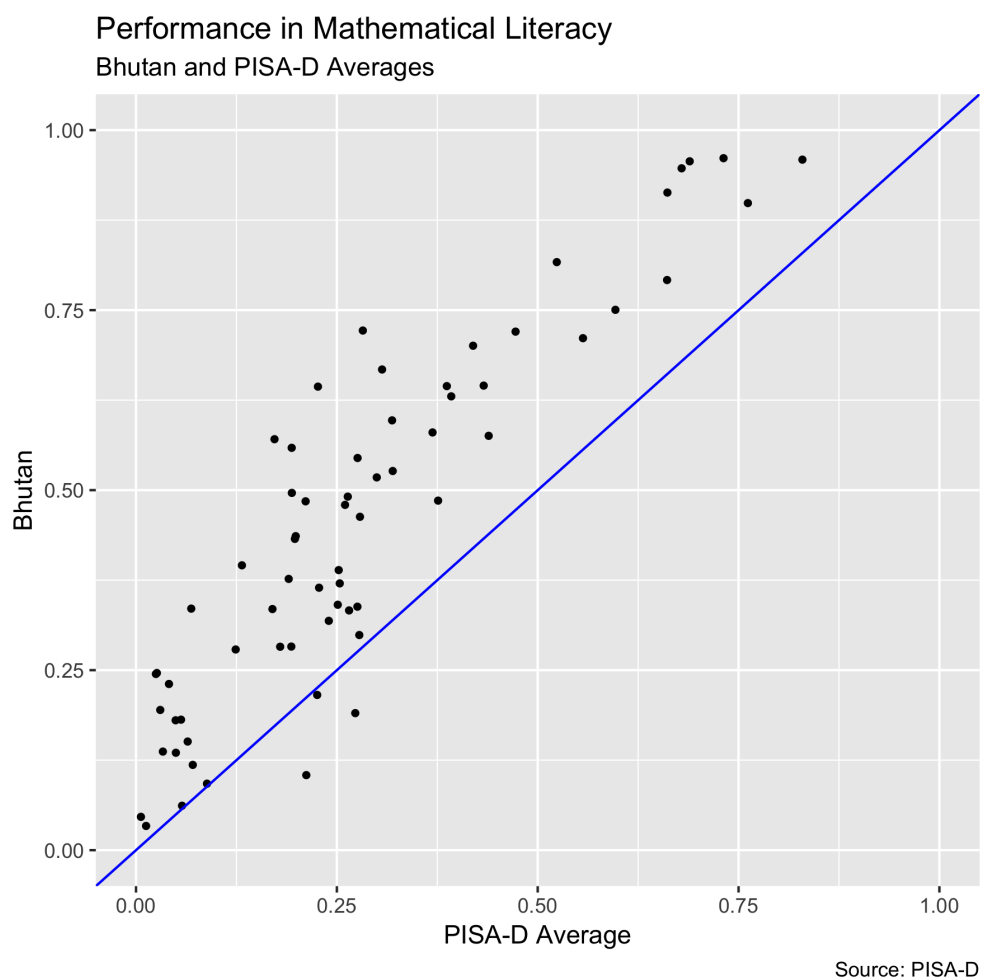
Bhutan’s students achieved an average solution rate of 38.84 percent in the PISA-D 2017 Mathematical literacy assessment, which was significantly higher than the PISA-D average solution rate of 30.65 percent.

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 compares the solution rates of students who took the PISA-D 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 blue diagonal line represents a benchmark situation in which Bhutan’s students achieved the same solution rate as students in comparison countries on average. The dots below the blue line are the items in which Bhutan’s students had lower solution rates than PISA-D average.

Figure 6: How students in Bhutan fared in PISA-D compared to the average of all PISA-D countries

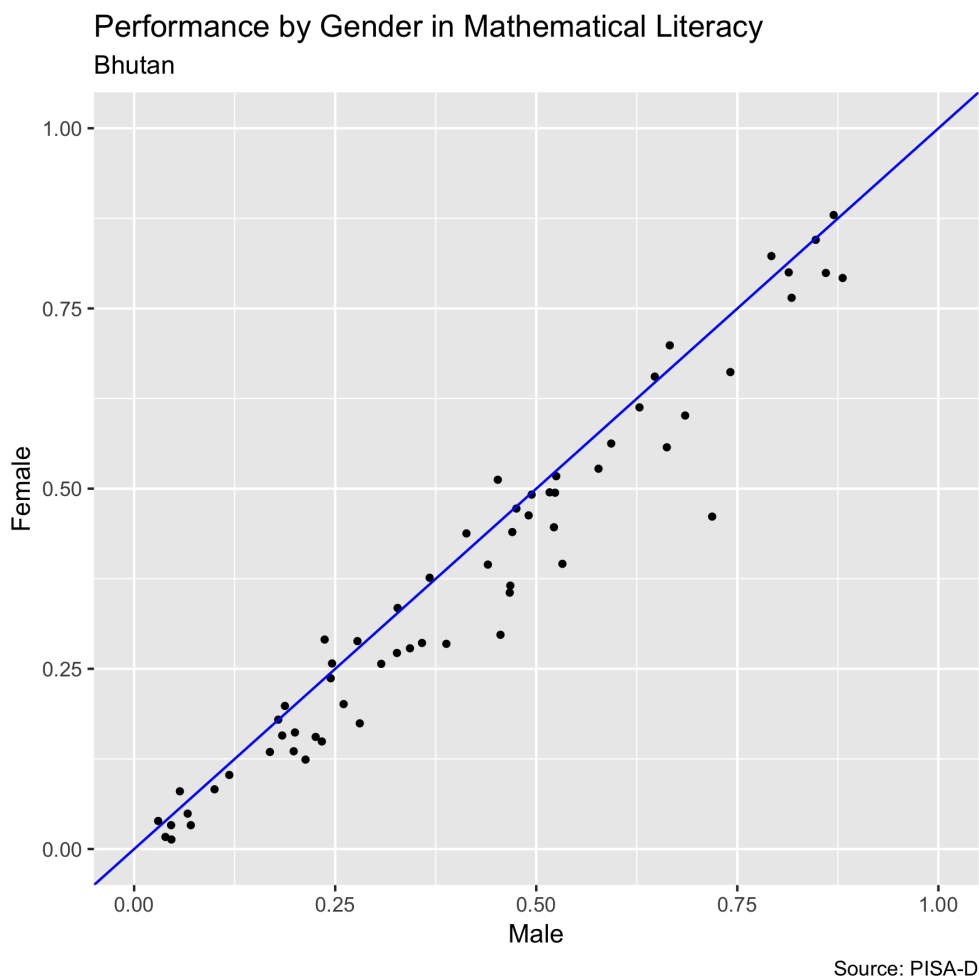


Percent-correct scatter plot, Bhutan vs PISA-D average (Mathematical literacy)

The visual inspection of Figure 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 Bhutan's 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-D test in Bhutan compares to the solution rate for girls on the same set of mathematics items. 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 literacy, 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

ITEM ID	ITEM NAME	BTN	DELTA PISA-D	DIFF	SITUATION / CONTEXT	CONTENT	PROCESS
pm915q02	Carbon Tax	12.2	15.6	-2.5	Societal	Change and Relationships	Employing Mathematical Concepts, Facts and Procedures
pm955q01	Migration	12.5	15.4	-1.9	Societal	Uncertainty and Data	Interpreting, Applying and Evaluating Mathematical Outcomes
pm5169q01	Shoe Sizes	13.5	16.3	-1.9	Personal	Change and Relationships	Employing Mathematical Concepts, Facts and Procedures
pm936q01	Seats in a Theatre	14.3	17.0	-1.8	Occupational	Change and Relationships	Employing Mathematical Concepts, Facts and Procedures



ITEM ID	ITEM NAME	DELTA			SITUATION / CONTEXT	CONTENT	PROCESS
		BTN	PISA-D	DIFF			
pm909q01	Speeding Fines	9.5	12.1	-1.7	Societal	Quantity	Interpreting, Applying and Evaluating Mathematical Outcomes
pm982q01	Employment Data	9.2	11.8	-1.6	Societal	Uncertainty and Data	Employing Mathematical Concepts, Facts and Procedures
pm650p001	Urban Population	13.0	15.4	-1.5	Societal	Uncertainty and Data	Interpreting, Applying and Evaluating Mathematical Outcomes

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. 4 of these items require employing mathematical concepts, facts and procedures; 3 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 7 are taken in Table 8 to compare percent correct of Bhutan with the percent correct of these items in the five reference countries, PISA-D and OECD averages.

Table 8: Relative strengths of Bhutan compared to the reference countries (PISA-D and PISA)

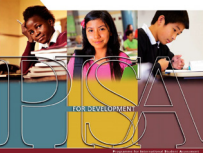
ITEM ID	ITEM NAME	BTN	ECU	PRY	PISA-D AVG	SGP	SWE	DOM	OECD AVG
pm915q02	Carbon Tax	57.63	27.52	20.04	25.47	85.71	50.12	8.66	64.17
pm955q01	Migration	55.07	42.09	26.01	27.84	NA	NA	NA	NA
pm5169q01	Shoe Sizes	45.35	36.91	11.22	20.57	NA	NA	NA	NA
pm936q01	Seats in a Theatre	37.22	27.08	7.21	15.58	NA	NA	8.71	30.78
pm909q01	Speeding Fines	80.88	76.02	68.52	58.48	92.88	84.97	44.25	84.17
pm982q01	Employment Data	82.70	72.71	64.66	61.99	87.00	86.07	51.98	84.97
pm650p001	Urban Population	50.47	41.34	28.04	27.23	NA	NA	NA	NA

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

ITEM ID	ITEM NAME	DELTA			ITEM FORMAT	CONTENT	PROCESS
		BTN	PISA-D	DIFF			
pm496q02	Cash Withdrawal	16.6	14.4	3.1	Open response	Quantity	Employing Mathematical Concepts, Facts and Procedures
pm919q02	Fan Merchandise	19.3	17.9	2.4	Open response	Quantity	Formulating Situations Mathematically
pm909q03	Speeding Fines	20.7	19.7	2.0	Open response	Change and Relationships	Interpreting, Applying and Evaluating Mathematical Outcomes
pm604p505a	Gas Gauge	15.3	14.5	1.8	Human coded	Quantity	Employing Mathematical Concepts, Facts and Procedures
pm954q02	Medicine Doses	20.3	19.5	1.7	Open response	Change and Relationships	Employing Mathematical Concepts, Facts and Procedures



ITEM ID	ITEM NAME	DELTA			ITEM FORMAT	CONTENT	PROCESS
		BTN	PISA-D	DIFF			
pm446q01	Thermometer Cricket	16.0	15.2	1.7	Open response	Change and Relationships	Formulating Situations Mathematically
pm5104q01s	Baby Growth	16.8	16.3	1.5	MCQ	Uncertainty and Data	Formulating Situations Mathematically
pm645p001	Airport Timetable	16.9	16.4	1.5	Human coded	Quantity	Interpreting, Applying and Evaluating Mathematical Outcomes

Table 9 displays 8 conspicuous items relating to differences between the delta values of Bhutan and the average of all PISA-D countries. 3 of these weak items require employing mathematical concepts, facts and procedures; 3 require formulating situations mathematically; and 2 require interpreting, applying and evaluating mathematical outcomes. 4 of the relative weaknesses are from the content area “Quantity”.

To classify the performance of Bhutan in more detail, the items identified in Table 9 are taken in Table 10 to compare percent correct of Bhutan with the percent correct of these items in five reference countries and PISA-D and OECD averages.

Table 10: Relative weaknesses of Bhutan compared to the reference countries (PISA-D and PISA)

ITEM ID	ITEM NAME	BTN	ECU	PRY	PISA-D AVG	SGP	SWE	DOM	OECD AVG
pm496q02	Cash Withdrawal	18.70	56.41	45.25	36.40	NA	NA	NA	NA
pm919q02	Fan Merchandise	5.71	22.52	12.68	11.00	NA	NA	NA	NA
pm909q03	Speeding Fines	2.70	7.52	4.57	4.71	NA	NA	NA	NA
pm604p505a	Gas Gauge	28.23	52.29	44.03	35.26	NA	NA	NA	NA
pm954q02	Medicine Doses	3.48	12.27	6.34	5.27	NA	NA	NA	NA
pm446q01	Thermometer Cricket	22.82	44.33	39.44	28.76	NA	NA	NA	NA
pm5104q01s	Baby Growth	16.97	38.92	23.39	20.56	NA	NA	NA	NA
pm645p001	Airport Timetable	16.38	32.42	25.16	19.74	NA	NA	NA	NA

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 content area “Quality”.

The delta values of items in strengths area ranged from 9.2 to 14.3 and delta values in weaknesses area ranged from 15.3 to 20.7. It is clear that Bhutan’s students found mathematics questions generally difficult.

Table 11 shows that of the three processes, Bhutan’s students are good in interpreting, applying and evaluating mathematical outcomes compared to most of the PISA-D countries.

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

ITEM GROUP	BTN	ECU	GTM	HND	KHM	PRY	SEN	ZMB	PISA-D AVG
Change and relationships	15.51	15.09	17.03	16.24	16.43	16.92	16.67	18.54	16.70
Quantity	14.47	13.81	15.17	14.89	15.53	15.31	15.59	17.39	15.38
Space and shape	13.98	13.71	14.95	14.88	14.82	15.01	15.30	16.64	15.04
Uncertainty and data	14.02	13.96	15.07	15.01	15.44	14.93	15.49	16.52	15.20
Occupational	13.64	13.51	15.32	14.67	14.89	15.76	14.80	16.87	15.12
Personal	14.22	13.55	15.09	14.67	15.12	15.15	15.59	16.66	15.12
Scientific	15.21	14.41	15.92	15.91	16.05	15.60	16.22	18.30	16.06

ITEM GROUP	BTN	ECU	GTM	HND	KHM	PRY	SEN	ZMB	PISA-D AVG
Societal	14.63	14.39	15.46	15.23	15.86	15.43	16.00	17.24	15.66
Employing mathematical concepts, facts, procedures and reasoning	14.26	14.18	15.69	15.10	15.50	15.68	15.67	16.99	15.55
Formulating situations mathematically	16.94	15.76	17.41	17.25	17.71	17.44	17.17	19.50	17.46
Interpreting, applying and evaluating mathematical outcomes	13.34	13.03	14.14	14.14	14.52	14.18	15.03	16.51	14.51
Selected-response	13.65	13.54	14.43	14.39	14.64	14.25	14.94	15.42	14.51
Open and close constructed-response	14.98	14.37	16.05	15.61	16.10	16.18	16.17	18.34	16.12

Analysis of illustrating items

In this section, we will illustrate some of the patterns observed in Table 7 and 9 with a typical PISA released items. 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 strength 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

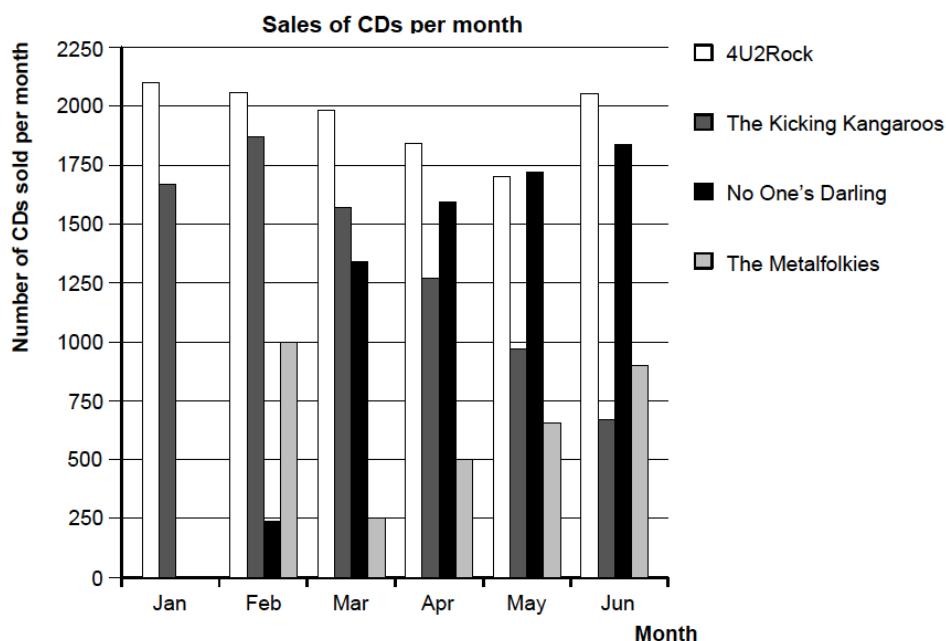
This is a simple multiple-choice item that requires comparison of speed when travelling 4 km in 10 minutes versus 2 km in 5 minutes. It is been classified within the employing process category because it requires the precise mathematical understanding that speed

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).

Content	Change and relationship
Process	Employing Mathematical Concepts, Facts and Procedures
Item format	Simple multiple choice

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: _____

Source: OECD (2013b), "PISA 2012 Released Mathematics Items", May 2013, www.oecd.org/pisa/pisaproducts/pisa2012-2006-relitems-maths-ENG.pdf.

Content	Uncertainty and data
Process	Formulating Situations Mathematically
Item format	Simple multiple choice

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 | D. 7,100 |
| B. 710 | E. 7,400 |
| C. 3,400 | |

Ans: _____

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

Content	Uncertainty and data
Process	Formulating Situations Mathematically
Item format	Simple multiple choice

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

Content	Change and relationship
Process	Formulating Situations Mathematically
Item format	Open constructed response



Scientific literacy assessment in PISA-D

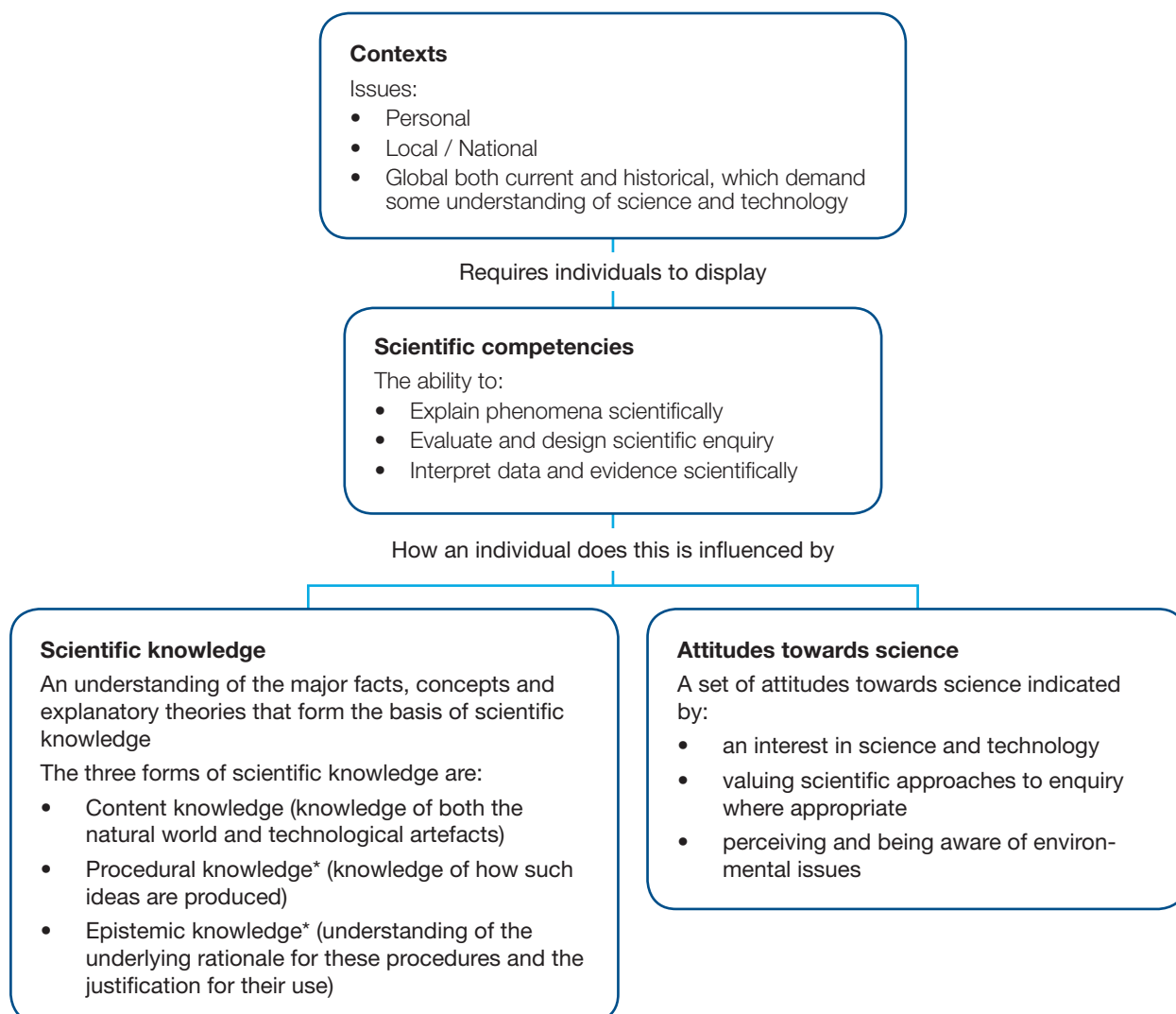
Scientific literacy is defined as the ability to engage with science-related issues, and with the ideas of science, as a reflective citizen. A scientifically literate person is willing to engage in reasoned discourse about science and technology which requires the competencies to explain phenomena scientifically,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

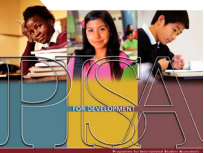


The characteristics of the science domain assessed are shown in Figure 8.

Figure 8: Aspects of the Scientific literacy assessment framework



*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

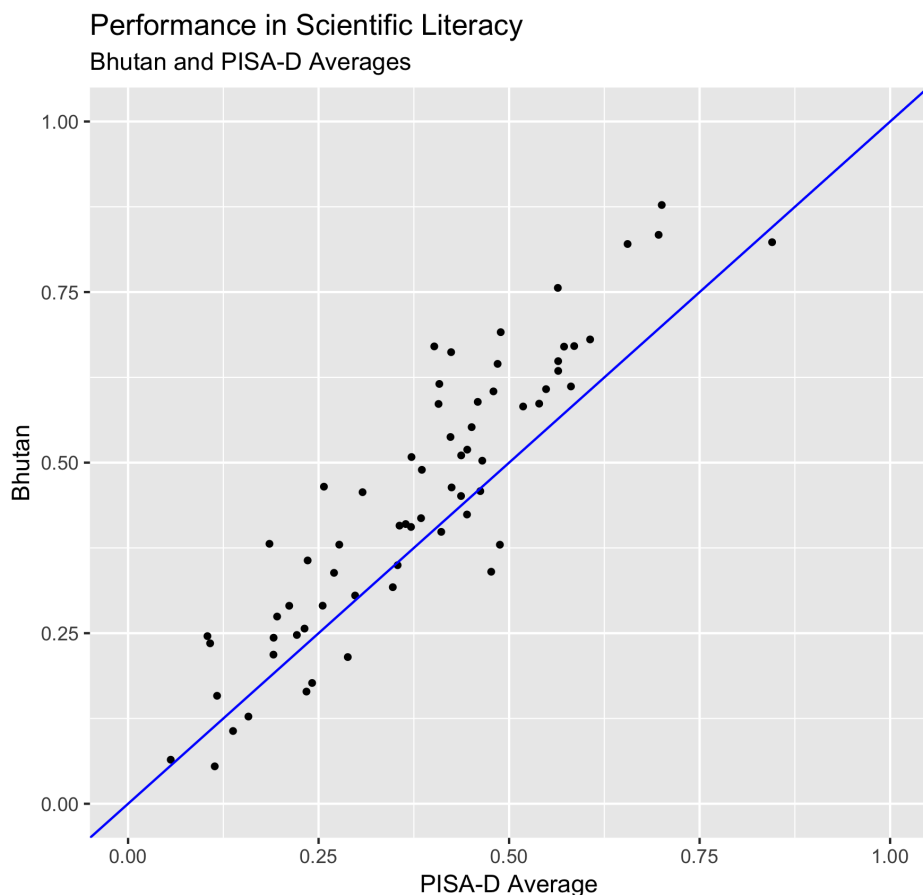
Bhutan's students achieved an average solution rate of 45.10 percent in the PISA-D 2017 Scientific literacy assessment, which was significantly higher than the PISA-D average solution rate of 38.28 percent.

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 rates of students who took the PISA-D 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 achieved the same solution rate as students in comparison countries on average. The dots below the blue line are the items in which Bhutan's students had lower solution rates than PISA-D average.

Figure 9: How students in Bhutan fared in PISA-D compared to the average of all PISA-D countries



Percent-correct
scatter plot, Bhutan
vs PISA-D average
(Scientific literacy)



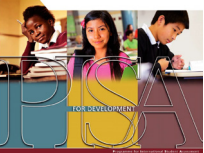
The visual inspection of Figure 9 clearly shows that on the majority items, students in Bhutan had higher success rates compared to PISA-D averages. The fact that not all items are perfectly aligned, and that, for a given success rate (and hence, difficulty level) in the comparison countries, there is 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 Bhutan's 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-D test in Bhutan compares to the solution rate for girls on the same set of science items. 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



Percent-correct scatter plot, Boys vs Girls



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

ITEM ID	ITEM NAME	BTN	DELTA PISA-D	DIFF	ITEM FORMAT	KNOWLEDGE	SYSTEM	COMPETENCY
ps476q01s	Heart Surgery	11.24	14.00	-1.98	MCQ	Content	Living	Explain phenomena scientifically
ps269q03	Earth's Temperature	15.75	18.34	-1.80	Open response	Content	Living	Explain phenomena scientifically
ps498q04	Experimental Digestion	14.21	16.73	-1.73	Open response	Procedural	Living	Interpret data and evidence scientifically
ps527q03as	Extinction of the Dinosaurs	11.33	13.77	-1.65	MCQ	Content	Earth and Space	Explain phenomena scientifically
ps256q01s	Spoons	8.35	10.77	-1.63	MCQ	Content	Physical	Explain phenomena scientifically
ps437q03s	Extinguishing Fires	13.35	15.67	-1.53	MCQ	Content	Physical	Explain phenomena scientifically

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 Bhutan's students can be seen primarily in items that require competency in explaining phenomena scientifically using knowledge "Content".

To classify the performance of Bhutan in more detail, the items identified in Table 12 are taken in Table 13 to compare percent correct of Bhutan with the percent correct of these items in the five reference countries and PISA-D and OECD averages.

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

ITEM ID	ITEM NAME	BTN	ECU	PRY	PISA-D AVG	SGP	SWE	DOM	OECD AVG
ps476q01s	Heart Surgery	67.04	51.50	43.02	40.18	84.65	64.91	34.89	68.51
ps269q03	Earth's Temperature	24.57	21.34	12.47	10.42				NA
ps498q04	Experimental Digestion	38.10	31.51	15.77	18.55				NA
ps527q03as	Extinction of the Dinosaurs	66.19	46.81	40.24	42.39				NA
ps256q01s	Spoons	87.76	72.06	74.17	70.03	98.33	88.42	68.57	88.81
ps437q03s	Extinguishing Fires	46.48	40.99	26.89	25.70	57.54	49.32	23.92	49.40

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.



Relative weaknesses

Table 14: Relative weaknesses of Bhutan compared to the average of all PISA-D countries

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

Table 14 displays 7 conspicuous items relating to differences between the delta values of the average of all PISA-D countries and Bhutan. The weakness of the Bhutan's students can be seen primarily in items that require interpreting data and evidence scientifically and evaluating and designing scientific enquiry. Most of these items required students' good grasp in content knowledge from "Physical" and "Earth and space" systems.

To classify the performance of Bhutan in more detail, the items identified in Table 14 are taken in Table 15 to compare percent correct of Bhutan with the percent correct of these items in five reference countries and PISA-D and OECD averages.

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

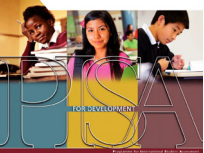
ITEM ID	ITEM NAME	BTN	ECU	PRY	PISA-D AVG	SGP	SWE	DOM	OECD AVG
ps437q06	Extinguishing Fires	21.50	39.35	34.66	28.83				NA
ps252q01s	South Rainea	17.71	31.81	33.22	24.15	53.91	58.02	18.48	52.94
ps408q05s	Wild Oat Grass	16.44	30.58	24.52	23.41	44.40	41.49	12.92	39.17
ps415q07s	Solar Panels	37.97	59.83	54.59	48.79	67.52	69.06	36.42	73.61
ps7221q1as	Clean Drinking Water	82.31	97.97	93.28	84.51				NA
ps438q03	Green Parks	5.48	13.36	9.66	11.37				NA
ps413q05s	Plastic Age	34.01	58.13	56.01	47.65	66.91	60.53	37.26	67.58

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

ITEM GROUP	BTN	ECU	GTM	HND	KHM	PRY	SEN	ZMB	PISA-D AVG
Evaluate and design scientific enquiry	14.28	13.52	14.58	14.48	14.83	14.49	15.11	15.26	14.61
Explain phenomena scientifically	13.36	13.31	14.07	13.90	14.75	13.91	14.75	15.14	14.26



ITEM GROUP	BTN	ECU	GTM	HND	KHM	PRY	SEN	ZMB	PISA-D AVG
Interpret data and evidence scientifically	13.54	13.44	14.30	14.16	14.90	13.97	14.79	15.26	14.40
Content knowledge	13.38	13.31	14.08	13.94	14.66	13.88	14.67	15.07	14.23
Epistemic knowledge	14.69	13.89	15.14	14.81	15.21	14.68	15.20	15.65	14.94
Procedural knowledge	13.57	13.35	14.19	14.11	14.98	14.13	15.07	15.32	14.45
Earth and space systems	13.33	13.17	13.94	13.91	14.68	13.62	14.59	14.95	14.12
Living systems	13.72	13.58	14.48	14.33	15.32	14.30	15.15	15.39	14.65
Physical systems	13.62	13.30	14.14	13.90	14.25	14.03	14.61	15.14	14.20
High cognitive demand	15.62	14.63	15.09	15.03	15.76	15.09	15.61	16.52	15.39
Medium cognitive demand	13.73	13.41	14.41	14.28	14.82	14.09	14.85	15.22	14.44
Low cognitive demand	13.15	13.19	13.91	13.74	14.66	13.84	14.70	14.99	14.15
Multiple choice items	13.36	13.16	13.93	13.87	14.52	13.79	14.72	14.78	14.11
Constructed-response items	15.02	14.81	16.15	15.48	16.59	15.65	15.55	17.86	16.01

Analysis of illustrating items

In this section, we will illustrate some of the patterns observed in Table 12 and 14 with a typical PISA released items. 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 strength 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

PISA 2015

Bee Colony Collapse Disorder
Question 4 / 5

Refer to “Exposure to Imidacloprid” on the right. Type your answer to the question.

Look at the result in week 20 for the hives that the researchers did not expose to imidacloprid (0 $\mu\text{g/kg}$). What does it indicate about causes of collapse among the studied colonies?

BEE COLONY COLLAPSE DISORDER
Exposure to Imidacloprid

Scientists believe that there are multiple causes for colony collapse disorder. One possible cause is the insecticide imidacloprid, which may cause bees to lose their sense of orientation when outside the hive.

Researchers tested whether exposure to imidacloprid leads to colony collapse. In a number of hives, they added the insecticide to the bees’ food for three weeks. Different hives were exposed to different concentrations of the insecticide, measured in micrograms of insecticide per kilogram of food ($\mu\text{g/kg}$). Some hives were not exposed to any insecticide.

None of the colonies collapsed immediately after exposure to the insecticide. However, by week 14, some of the hives had been abandoned. The following graph records the observed results:

Number of Weeks After Exposure to Insecticide	0 $\mu\text{g/kg}$	20 $\mu\text{g/kg}$	400 $\mu\text{g/kg}$
10	0%	0%	0%
12	0%	0%	0%
14	0%	25%	50%
16	0%	25%	50%
18	0%	25%	100%
20	25%	75%	100%
22	25%	100%	100%

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.

Competency

Explain Phenomena Scientifically

Knowledge – System

Content – Living

Context	Local/National – Environmental Quality
Cognitive Demand	Medium
Item Format	Open Response – Human Coded

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

PISA 2015

?

◀

▶

Slope-Face Investigation

Question 1 / 2


Refer to "Data Collection" on the right. Type your answer to the question.

In investigating the difference in vegetation from one slope to the other, why did the students place two of each instrument on each slope?


SLOPE-FACE INVESTIGATION

Data Collection


The students place two of each of the following three instruments on each slope, as shown below.



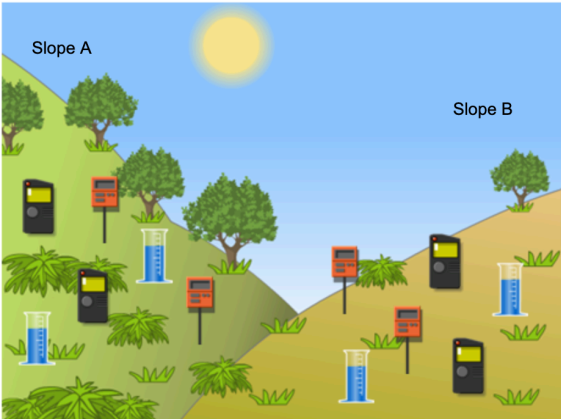
Solar radiation sensor: measures the amount of sunlight, in megajoules per square metre (MJ/m²)



Soil moisture sensor: measures the amount of water as a percentage of a volume of soil



Rain gauge: measures the amount of rainfall, in millimetres (mm)



Source: 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

PISA 2015

Running in Hot Weather
Question 4 / 6

► **How to Run the Simulation**

Run the simulation to collect data based on the information below. Click on a choice, select data in the table, and then type an explanation to answer the question.

Based on the simulation, when the air humidity is 40%, what is the highest air temperature at which a person can run for one hour without getting heat stroke?

☐ 20°C
☐ 25°C
☐ 30°C
☐ 35°C
☐ 40°C

★ Select two rows of data in the table to support your answer.

Explain how this data supports your answer.

Air Temperature (°C)
 Air Humidity (%)
 Drinking Water ☒ Yes ☐ No

Run

Air Temperature (°C)	Air Humidity (%)	Drinking Water	Sweat Volume (Litres)	Water Loss (%)	Body Temperature (°C)

Simulation Visuals:

- Sweat Volume (Litres):** A vertical scale from 0 to 3. The level is currently at approximately 2.5.
- Water Loss (%):** A vertical scale from 0 to 5. The level is currently at approximately 2.5, labeled "Dehydration".
- Body Temperature (°C):** A vertical scale from 36 to 42. The level is currently at approximately 40.5, labeled "Heat Stroke".

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.

Competency	Evaluate and Design Scientific Enquiry
Knowledge – System	Procedural
Context	Personal – Health and Disease
Cognitive Demand	Medium (Level 4)
Item Format	Open Response – Human Coded

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 support 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 system still faces 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 future cycle of PISA.

Summary of findings

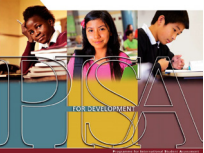
Overall performance

Bhutan's performance in PISA-D is remarkably similar to low and middle-income countries. Although no PISA scores were generated for Bhutan due to small sample size, a reliable estimate based on the percent correct scores shows students' performance significantly below OECD average. When compared to other PISA-D countries, the performance of Bhutan's students lies in between the two highest performing PISA-D countries (Ecuador and Paraguay).

Summary of relative strengths in the three domains

In Reading literacy, Bhutan's students had significantly 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, Bhutan's 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, Bhutan's 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 judgement drawing on standards beyond the text. Much of these texts are associated with out-of-school settings for students, such as rules for clubs and records of games, which often take place unofficially at schools. Such items are found more difficult as they require 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 require 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 require students to evaluate reports of scientific findings and investigations critically. These tasks rely 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.

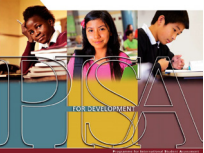
The following section highlights the key messages based on PISA-D general findings which have implications on students' performance:

- 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 National Education Assessment;
- Bhutan Council for School Examinations and Assessment should enhance 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 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 improve the student learning outcomes; and
- Parents and teachers should break the gender stereotypes about science / mathematics / reading related activities and occupations to allow children achieve their full potentials.

Key messages based on PISA-D general findings which have implications on students' performance in the three domains are mentioned below.

Reading literacy

- Schools should institute reading culture; strengthen teacher model reading; offer literary activities and provide varieties of reading materials to enhance students' 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.

Mathematical literacy

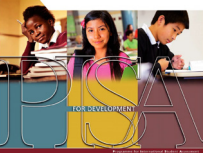
- Mathematics teachers should use diverse teaching strategies considering both the content and the learners with different abilities, motivation and interests. The focus should be on student-centered teaching rather than teacher-centered teaching.
- Apart from teaching fundamental elements of the mathematics 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 provide ICT facilities, better Internet connectivity and promote integration of ICT in teaching-learning process.
- 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.

Next steps in Bhutan's participation in future cycles of PISA

Given the nature of Bhutan's participation in PISA-D, in-depth understanding of the nature of strengths and weaknesses cannot be ascertained on a significant level to be able to recommend focused measures and strategies for improvement. However with Bhutan's participation in future PISA cycles, 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

1. ACER. (2017). *TIMSS 2015 reporting Australia's results*. Victoria, Australia: ACER Press.
2. BCSEA. (2015). *BCSEA journal for educational assessment*. Thimphu, Bhutan: BCSEA
3. HO, E. S.-c. (2013). *Student learning assessment*. Bangkok: UNESCO.
4. Kagan, S. (2003). *A brief history of Kagan structures*. San Clemente, CA: Kagan Publishing.
5. Luo, E. C. (2016). *Assessment of transversal competencies: Policy and practice in the Asia-Pacific region*. Bangkok: UNESCO.
6. M.o.E. (n.d.). *Towards educating for GNH: A guide to school self-assessment tool and school improvement plan*. Thimphu, Bhutan: Education Monitoring and Support Service Division, MoE.
7. MoE. (2014). *Bhutan Education Blueprint (2014-2024)*. Thimphu, Bhutan: MoE.
8. OECD. (2000). *Measuring student knowledge and skills: The PISA 2000 assessment of reading, mathematical and scientific literacy*. Paris, France: OECD Publishing.
9. OECD. (2003). *Mathematics teaching and learning strategies*. Paris, France: OECD Publishing.
10. OECD. (2016). *PISA 2015 results (Volume I): Excellence and equity in education*. Paris, France: OECD Publishing.
11. OECD. (2016). *Ten questions for mathematics teachers and how PISA can help answer them*. Paris: OECD Publishing. DOI: <http://dx.doi.org/10.1787/9789264265387-en>
12. OECD. (2017). *PISA for development assessment and analytical framework: Reading, mathematics and science, preliminary version*. Paris, France: OECD Publishing.
13. OECD. (2018). *Effective teacher policies: Insights from PISA*. Paris, France: OECD Publishing. DOI: <http://dx.doi.org/10.1787/9789264301603-en>
14. Schleicher, A. (2018). *World class: How to build a 21st-century school system*. Paris, France: OECD Publishing.

“An educated and enlightened society of Gross National Happiness, built and sustained on the unique Bhutanese values of **THA DAM-TSIG LEY GJU-DREY**”

┌ **Vision** - Ministry of Education, Bhutan ┐

“An internationally recognized centre of excellence in educational assessment that provides quality services to build the integrity and profile of the education system leading to an improvement in the quality of learning in Bhutan”

┌ **Vision** - Bhutan Council for School Examinations and Assessment, Bhutan ┐