

RESEARCH

Open Access



Learning data science in elementary school mathematics: a comparative curriculum analysis

Yook Kit Ow-Yeong, Ibrahim H. Yeter and Farhan Ali*

Abstract

Background Data literacy is increasingly important in today's data-driven world. Students across many educational systems first formally learn about data in elementary school not as a separate subject but via the mathematics curriculum. This experience can create tensions in the priorities of learning and assessment given the presence of other foundational mathematics content domains such as numbers, algebra, measurement, and geometry. There is a need to study data literacy in comparison to these other content domains in elementary mathematics. To address this need, we developed a methodology motivated by thinking curriculum theory and aligned with international assessment framework, for comparative analysis across mathematics content domains. This methodology examined increasing levels of cognitive domains from knowing to applying to reasoning across mathematics content domains. Intended, assessed, and attained curricula were analyzed using Singapore as a case study, combined with broader comparisons to attainments in four East Asian countries in TIMSS, an international large-scale assessment.

Results We found that learning in the data domain had very limited coverage in intended and assessed curricula in Singapore. However, compared to other mathematics content domains, the data curriculum placed heavier emphasis on higher-order cognitive domains including the use of generally difficult mixed data visualizations. This demanding curriculum in Singapore was associated with the highest attainment in the data domain among average 4th grade Singaporean students relative to students in four East Asian countries in TIMSS, as analyzed by quantile regression. However, lower-performing Singaporean students at the 10th percentile generally did not outperform their East Asian peers. We further found very limited applications of data in other mathematics domains or cross-domain learning more generally.

Conclusion Our study offers a comparative analysis of the data curriculum in elementary school mathematics education. While the data curriculum was cognitively demanding and translated to very high average attainments of Singaporean students, the curriculum did not equally help weaker Singaporean students, with implications on current discourse on equity–excellence trade-off in science, technology, engineering, and mathematics (STEM) education. Our study further highlights the lack of cross-domain learning in mathematics involving data. Despite the broad applicability of data science, elementary school students' first formal experience with data may lack emphasis on its cross-domain applications, suggesting a need to further integrate data skills and competencies into the mathematics curriculum and beyond.

Keywords Curriculum, Data science, Mathematics, Statistics, Singapore, East Asia, TIMSS

Introduction

As the world produces digital data at ever-increasing rates, there is a need to make sense of the large mass of data to generate new knowledge. Consequently, in the past few years, there has been tremendous interest in data science as an emerging field (Cao, 2017). Educators

*Correspondence:

Farhan Ali

farhan.ali@nie.edu.sg

National Institute of Education, Nanyang Technological University, 1
Nanyang Walk, Singapore 637616, Singapore

have recently called for a more prominent K-12 data science education in order to foster early data literacy and support the increased demand for data literate citizens (Wise, 2020). Bakker et al. (2021) recently surveyed mathematics educational researchers, reporting data literacy as among the most frequently mentioned future educational goals of mathematics. The most recent Guidelines for Assessment and Instruction in Statistics Education II (GAISE II), a comprehensive document guiding K-12 statistics and data science education, highlight the goal of developing students as data problem-solvers (Bargagliotti et al., 2020, 2021). Given that learning about data is traditionally embedded within elementary mathematics curricula in many countries (Davies & Sheldon, 2021; Groth, 2018), in this study, we aim to understand how and to what extent data knowledge and skills are learned and assessed within existing mathematics curriculum in comparison to other mathematics content domains. We focus on Singapore as a case study while further using multi-country attained curriculum data from international assessment for broader comparisons.

Literature review

In the ensuing literature review, we draw on two major theoretical frameworks and related research. One examines the relationship between data science, statistics and mathematics, and the other focuses on curriculum framework for a comparative analysis. Both lay the ground for research in comparing how students learn about data and are assessed in relation to other fundamental content domains in elementary mathematics education.

Learning about data in elementary mathematics

Data is¹ a group of numbers in specific contexts, giving them meaning beyond their abstract representation (Cobb & Moore, 1997). The study of data has disciplinary roots in statistics (Donoho, 2017). However, there is consensus that the emerging field of data science is an amalgamation of multiple disciplines beyond statistics, for a number of reasons (Blei & Smyth, 2017). First, data deluge has necessitated computational methods to manage, process and mine big data (Cao, 2017). Moreover, there is a shift toward working with messy real-world data in different domains, giving rise to important skills in data processing (e.g., cleaning, transformation) and data visualization, going beyond the traditional focus of theory-driven statistical methods (Donoho, 2017). Similar to previous expositions (Cao, 2017; Engel, 2017), we consider data science as an interdisciplinary science of learning from data, aided by modern computational tools

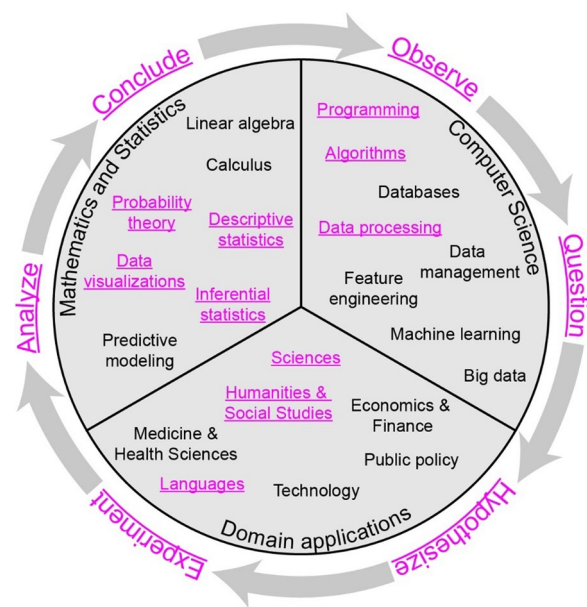


Fig. 1 Data science as an interdisciplinary field. Our framework of data science as an amalgamation of mathematics, statistics, and computer science with specific domain applications. Overlaid are iterative scientific processes that support data science. Underlined in magenta, components of data science formally incorporated in the curriculum of many elementary and secondary school systems

and methods. Data science combines the disciplines of statistics, mathematics, and computer science, applied to specific domains (Fig. 1). Further extending previous expositions, we also make explicit the scientific nature of data science. First, processes that classically make up the scientific methodology are important to data science such as formulating hypotheses and testing them using various forms of data collection and computational experiments (Karpatne et al., 2017). Second, key elements of how the scientific endeavor operates also play central roles in data science such as cycling through data for discoveries and reliance on reproducibility (Blei & Smyth, 2017). Some components of data science, as thus conceptualized, are already formally incorporated in the curriculum of elementary and secondary school systems via data visualizations (e.g., graph learning) (Aksoy & Bostan, 2021), descriptive and inferential statistics (e.g., distributional reasoning) (Biehler et al., 2018), data processing (e.g., in science inquiry) (Watson, 2017), and programming (Bråting & Kilhamn, 2021).

As an important component of data science (Fig. 1), statistics is typically incorporated in K-12 education via the subject of mathematics (Groth, 2018). This situation motivates a comparative approach to statistics education within mathematics. Such studies have employed curriculum analysis, for example, qualitatively comparing the intended mathematics curriculum across different

¹ Used here in the singular form, etymology notwithstanding.

content domains such as numbers, algebra, geometry, and measurement (Dingman et al., 2013; Lv & Cao, 2018). Davies and Sheldon (2021) comprehensively reviewed the curricular challenges of embedding data science and statistics in England's national mathematics curriculum. For example, standard paper-based assessment in mathematics is argued to be inappropriate for statistics as the latter requires experience working with actual data, most meaningfully done on a computer, unlike for traditional mathematics contents (Davies & Sheldon, 2021). In addition, Jones et al. (2015) performed an analysis of mathematics textbook curricula in the United States, finding that statistics contents tended to be very limited in earlier grades but gradually increased to about 20% of instructional pages by 5th grade. Tellingly, the textbooks at most grade levels covered statistics later compared to other mathematics contents, likely reflecting (intentionally or unintentionally) the order of importance assigned by the textbook writers (Jones et al., 2015). There was also very little integration of statistics with other mathematics contents in all textbooks examined except one (Jones et al., 2015). Other studies have focused on teachers, with preservice mathematics teachers reporting the least confidence in teaching statistics compared to other mathematics content domains (Lovett & Lee, 2017). Indeed, objective assessments revealed inadequate preservice teachers' foundational knowledge in statistics, especially at higher levels of statistical inquiry (Lovett & Lee, 2017). Shin (2021) further found that preservice teachers drew largely on mathematical pedagogical knowledge unrelated to statistical thinking when noticing statistics classroom interactions, suggesting an imbalance in their training of content-specific pedagogical knowledge. In short, students may not fully experience the unique rigors of thinking about data when learning within a mathematics curriculum. While these studies have contributed to a comparative understanding of statistics education within mathematics, key gaps remain. These include how and to what extent data skills and knowledge are learned and assessed within an existing mathematics curriculum in comparison to other mathematics content domains that seemingly play larger roles in elementary mathematics.

Potentially contributing to some of these gaps in knowledge is statistics educational research tending to mature separately from mathematics educational research. For example, two recent influential books on statistics education had very limited comparisons to other mathematics content domains (Ben-Zvi et al., 2017; Leavy et al., 2018). The history of statistics has also been traced not to mathematics but to demography and epidemiology in the seventeenth century followed by data visualizations in the eighteenth century (Wild et al., 2018). It also appears that research on statistics education seemed to

be infrequently published in leading mathematics education journals in the last decade, an observation that others have made (Batanero et al., 2011). Further, statistical thinking is argued to be distinct from mathematical thinking; chance and uncertainty play central roles in statistical thinking unlike in mathematical thinking (Cobb & Moore, 1997). All these observations suggest statistics education as aspiring to, and to some degree, succeeding in, growing separately from mathematics education. However, we argue for study of learning about data within the mathematics curriculum. This need is due to the strong entrenchment of statistics within the K-12 mathematics curriculum, likely to continue in the foreseeable future (Davies & Sheldon, 2021; Groth, 2018). Teachers trained in mathematics are the ones creating learning environments to formally teach children about data for the first time, with potential issues that may arise as reviewed above. Overall, there is a need to better understand the curriculum priorities, learning expectations, assessment, and attainment in statistics in comparison to other mathematics content domains.

Curriculum analysis framework

One approach to comparing across mathematics content domains is to use a curriculum framework. There are contested views of curriculum as being focused on content, product or process (Kelly, 2009). Here, following the suggestion of Hirsch and Rey (2009), curriculum is defined as what "society values and expects in terms of mathematics content" (p. 749). This definition makes clear that a curriculum is a society-specific statement of priorities, emphases, and expectations in mathematics. Curriculum can be divided into the explicit and hidden curriculum. While the latter is an important aspect of classroom and school experience (Alsubaie, 2015), the explicit curriculum lends itself to be more transparently quantified and compared within and across educational systems. Explicit curriculum can be further decomposed to three components most relevant to the present study: (1) the intended curriculum that officially documents the progression of learning contents and experiences; (2) the assessed curriculum that monitors student learning; and (3) the attained curriculum that quantifies actual extent of achievement of knowledge and skills. Analysis using these different components of curriculum is very useful as they provide a common basis with which to compare across mathematics content domains. This framework is also in line with studies cited above that have taken a curriculum approach involving various components such as intended curriculum and textbook curriculum.

Beyond the structural components of curriculum as outlined, there is also a process component of curriculum. Nisbet (2005) proposes the thinking curriculum in

which processes related to thinking are taught as cognitive skills. This skills-based approach of how students perceive, organize and make sense of information has origins in cognitive psychology but with alignment to educational frameworks including work by Bloom on hierarchy of learning objectives in the cognitive domain (Krathwohl, 2002). At the lowest level, students learn to reproduce knowledge, typically involving memorization and regurgitation of isolated facts or procedures, processes that educational systems have traditionally focused on (Nisbet, 2005). Other cognitive skills are higher order such as applying existing knowledge to solve a problem or even higher, reasoning based on multiple pieces of information, a skill that is championed as part of education in the 21st Century (Geisinger, 2016). This cognitive skills approach is general enough to be applied to different content domains within a discipline, and is used by different curriculum frameworks in both local (Ministry of Education Singapore, 2012; Tanudjaya & Doorman, 2020) and international assessments (Lindquist et al., 2017).

Curriculum is embedded within a specific context of values and expectations as argued by Hirsch and Rey (2009) and thus, needs to be studied as such. The Singapore curriculum can be used as a case study for several reasons. Singapore is a small, developed Southeast Asian country with an educational system that is generally well regarded based on international assessments (Mullis et al., 2020; Schleicher, 2019). However, past research has highlighted existence of gaps in educational achievement, some of which are wide (Ali, 2016), suggesting a more complex picture that deserves further investigation. Singapore is also a good case study as the local literature on statistics education is scant (Chia, 2016; Wu & Wong, 2007), with mostly qualitative approaches. None, to our knowledge, has taken a curricular perspective for comparative purposes. This gap is not surprising as statistics, like in many countries, was only recently introduced in the past 20 years in Singapore compared to much longer history of educational research on traditional mathematics content domains (Toh et al., 2019). While our main intended and assessed curricular analyses focus on Singapore, a broader international comparison is useful. This comparison can quantify the extent to which the local intended and assessed curricula translate to actual attainments using common yardsticks across countries. For this cross-country comparison, international large-scale assessment data can be leveraged, specifically Trends in International Mathematics and Science Study (TIMSS). Large-scale assessments do have limitations such as questions about measurement and validity (Johansson, 2020). However, they can still be valuable in providing quality data to compare the attainments of Singaporean

students to similarly developed East Asian countries. The TIMSS data is also based on a well-established assessed curriculum framework with a hierarchy of cognitive skills aligned to what is reviewed above (Lindquist et al., 2017). We can further analyze the performance in specific content domains to disentangle country-specific effects (e.g., high attainments by a country regardless of content domains) from mathematics domain-specific patterns that apply across countries (e.g., lower attainment in a particular domain for all countries). There is also data on how well the local curriculum matches the TIMSS assessed curriculum to further explain the comparative results (Fishbein et al., 2021). Overall, a curriculum framework allows comparative analysis within Singapore across mathematics content and across countries for a more comprehensive picture of how students learn about data in mathematics.

Research objectives

Focusing on the mathematics curriculum in Singapore, our study has the following research objectives (RO) and associated research questions (RQ):

RO1) To compare the data domain to other mathematics content domains in the intended curriculum: RQ1a, What is the intended content coverage of the data domain? RQ1b, How does it compare to other content domains in terms of cognitive skills required?

RO2) To compare the data domain to other mathematics content domains in the assessed curriculum: RQ2a, What proportion of the assessed curriculum is devoted to the data domain compared to other content domains? RQ2b, Are there differences in the assessed cognitive skills required by the different content domains? RQ2c, To what extent is there cross-content domain learning, especially involving data, and if so, how is it being assessed?

RO3) To compare Singaporean students' attainment using international large-scale assessment: RQ3a, How do the attainments of Singaporean students in the current 2019 cycle compare to Singaporean students in previous cycles? RQ3b How do the attainments of Singaporean students compare to East Asian peers in the current 2019 cycle? RQ3c How are these differences, if any, explained by differences in the local curriculum?

Methods

Intended curriculum

We examined the 2013 Singapore Primary School Mathematics Syllabus document written in English (Ministry of Education Singapore, 2012). The document forms the foundation of the intended elementary mathematics curriculum in Singapore's public school system (Lee et al., 2019), attended by the vast majority of

students in Singapore. The 2013 document is the most recent complete curriculum for all elementary grades. We programmatically extracted verbs from the ‘Learning Experiences’ (henceforth, learning verbs) section of each of the three content domains (termed strands in the official document): numbers and algebra, measurement and geometry, and statistics. This section provided detailed descriptions of learning experiences intended for students. Examples included “Write addition and subtraction equations for number stories” and “Use data from the Internet to make a picture graph”. The TIMSS 2019 mathematics framework (Lindquist et al., 2017) lists an array of verbs for three cognitive skills labeled as ‘cognitive domains’: knowing, applying, and reasoning. Descriptions of these domains, as given below, were used to classify the learning verbs in the intended curriculum. For verbs not in TIMSS definitions, authors discussed and assigned them to a cognitive domain based on how the word was used in the intended curriculum, supplemented by past reviews of verb usage in learning objectives (e.g., Stanny, 2016). For the statistics domain, we focused on the ‘Data representation and interpretation’ sub-domain which dominates the elementary school statistics curriculum. The other sub-domain is ‘Data analysis’, a much smaller content area involving average. In addition to being a small content area, ‘Data analysis’ is intended to be taught very late at Primary 5 (second to the last grade for elementary schools in Singapore). Because we were interested in comparing content sequence across primary (grade) levels, it made ‘data analysis’ sub-domain even less useful given the very late introduction. Thus, we narrowed our focus and simply labeled the remaining ‘Data representation and interpretation’ sub-domain as ‘data,’ which is also aligned with the TIMSS content domain labeling. We quantified the percentage of words belonging to each cognitive domain in each of the three content domains. Verb extraction, frequency counts in word clouds and bar graphs were generated using Python programming language.

Assessed curriculum

For assessed curriculum, we analyzed items from recent semestral assessment 2 (SA2) in public schools. These assessments were based on the 2013 intended curriculum as elaborated above. SA2 are summative assessments that Singaporean students take toward the end of the school year that would generally cover all of the school year’s contents, and are locally developed and administered in individual schools. We focused on Primary 2 to 6 as summative assessments are very rare at

Primary 1. Following other studies that have sampled schools in Singapore (e.g., Ang et al., 2020), we divided Singapore into three regions: eastern, central, and western. For each region, we randomly identified up to five public schools and attempted to obtain their assessment booklets across all levels (Primary 2–6). However, assessments for some schools were not made available and/or they may not have assessment for all levels. Eventually, six schools spread across the three regions contributed to our sample for complete assessment booklets across all levels. Two researchers independently categorized the assessment items to a cognitive domain—knowing, applying and reasoning—based on our definition that broadly aligned with TIMSS’ definitions. Generally, we defined knowing as testing a student’s knowledge using lower-order skills such as recalling and retrieving information. For example, this might involve doing two-digit addition for numbers domain or reading off a value from a bar graph for data domain. Applying entailed students utilizing knowledge in a range of situations involving intermediate order skills such as efficient problem-solving and data modeling. For example, students in the geometry domain were expected to identify and make use of properties of perpendicular or parallel lines to solve for angles in a complex figure while in the data domain, they were required to combine mathematical operations after reading data off a bar graph. Reasoning required students to think logically and systematically to synthesize novel ways to approach or solve problems with higher-order skills such as justifying solutions and drawing conclusions based on evidence. In the numbers domain, for example, students were expected to solve multi-step word problems requiring inference while in the data domain, students had to make and justify conclusions from one or more data displays. Our scheme of classifying items is congruent with past work on data literacy. For example, Curcio’s (1987) framework for graph comprehension has three increasingly complex skills. Our knowing classification approximately maps to Curcio’s “reading the data” while applying involves “reading between the data” and reasoning entails “reading beyond the data”. For assigning to cognitive domains, two raters had 88.9% agreement and discrepancies were resolved by consensus. The items were also assigned to content domains (numbers and algebra, measurement and geometry, and data) and a data visualization type (for data domain questions only). Items involving multiple domains were assigned as such. After categorization, all statistical analyses were done in Python.

Attained curriculum: cross-country analysis of TIMSS achievement data

We used publicly available TIMSS data from five cycles (2003, 2007, 2011, 2015, 2019) for Grade 4 mathematics.² The latest cycle in 2019 is of most interest as it came from the students who underwent the 2013 curriculum as elaborated above. The past cycles were further analyzed to address our research question of how attainments in the current cycle compared to past cycles over the decades. Curriculum changes in Singapore are known to be incremental instead of wholesale reforms as elaborated by a recent review (Lee et al., 2019). Thus, while attainments in previous cycles may have come from different curricula, there are overlaps in terms of knowledge and skills intended to be learned across the decade, supporting comparisons of results. In addition to Singapore, we also examined data from Hong Kong SAR, Republic of Korea, Chinese Taipei, and Japan. These countries have similar levels of socio-economic development, have comparable East Asian demographics as Singapore (over 70% of Singaporeans are of Chinese descent), and are all generally performative educational systems. Many studies in the past have also compared Singapore to these East Asian countries using TIMSS data or otherwise (e.g., Chen, 2014; Chen et al., 2018; Tan, 2018), thus, we used similar comparative analyses. TIMSS 2019 has three content domains: numbers, measurement and geometry, and data. The numbers domain also includes pre-algebra concepts involving computing unknown variables (Lindquist et al., 2017). These TIMSS content domains generally overlapped with Singapore's three content domains. Using total student weight (labeled as TOTWGT by TIMSS) and plausible values (PVs), we computed scale scores for Singapore and four East Asian countries in each of the three content domains. TIMSS analysis was done in R programming language using *intsvy*, a package for processing and analysis of large-scale assessment data given their unique sampling designs (Caro & Biecek, 2017).

To statistically compare attainments, we used quantile regression, a flexible statistical procedure that can examine specific locations in the distribution without assumptions of normality and linearity unlike in linear regression. In our case, we focused on the 50th percentile and 10th percentile, corresponding, respectively, to the middle- and lower-performing students that are of main interest in our study. Because there is a very large number of all possible pairwise tests that can be done (over 2000 possible pairwise tests from $5 \text{ (countries)} \times 5$

(cycles) $\times 3$ (content domains) number of values), we instead use planned comparisons derived specifically from our RQ3a and RQ3b (see text above). For RQ3a, we used quantile regression on the 50th percentile Singaporean students in the current 2019 cycle, comparing it to 50th percentile Singaporean students in the four previous cycles (2003, 2007, 2011, 2015). We performed this test for all three content domains. These tests were repeated for the 10th percentile. A total of 12 tests for 50th percentile and 12 tests for 10th percentile were conducted. We further analyzed cycle-on-cycle changes by comparing change from 2015 to 2019 with 2003 to 2007, 2007 to 2011 and 2011 to 2015 (unlike the above analysis which compared actual attainments in 2019 cycle to all other cycles). This analysis was done for both 50th and 10th percentile Singaporean students for the data domain only, resulting in 8 tests. For RQ3b, we used quantile regression on the 50th percentile Singaporean students in the current 2019 cycle, comparing it to 50th percentile of East Asian students in the current 2019 cycle. We performed this test for all three content domains. These tests were repeated for the 10th percentile performers. A total of 12 tests for 50th percentile and 12 tests for 10th percentile were conducted. Consistent with previous research, we performed separate tests for each PV, taking the average t-statistic across PVs as the final statistic to compute the p-values. We furthermore used total student weight (TOTWGT in TIMSS) as weights. These steps ensured that the tests incorporated uncertainty in estimating student performance (PV) as well as national representativeness (weights). Even though the comparisons were planned, there was still a large number of comparisons. To be more conservative, we used Bonferroni correction to maintain family-wise Type 1 error at 0.05. P-values reported have been corrected for multiple planned comparisons. Furthermore, because of the large sample sizes, most comparisons were highly statistically significant. Effect sizes (Cohen's d) were thus further computed to aid interpretation of statistically significant comparisons. Statistical analyses were done on SPSS.

Curriculum overlap was analyzed for Singapore and the East Asian countries using the Test-Curriculum Matching Analysis (TCMA) data that reported whether TIMSS assessment items were covered in the national curriculum as determined by experts in individual countries (Fishbein et al., 2021). Using the TIMSS database, we assigned the test items to their respective content domains, allowing us to examine whether Singapore's performance was due to differences in test item coverage.

² Equivalent to Primary 4 in Singapore. The term Grade 4 is kept to be consistent with TIMSS terminology where appropriate.

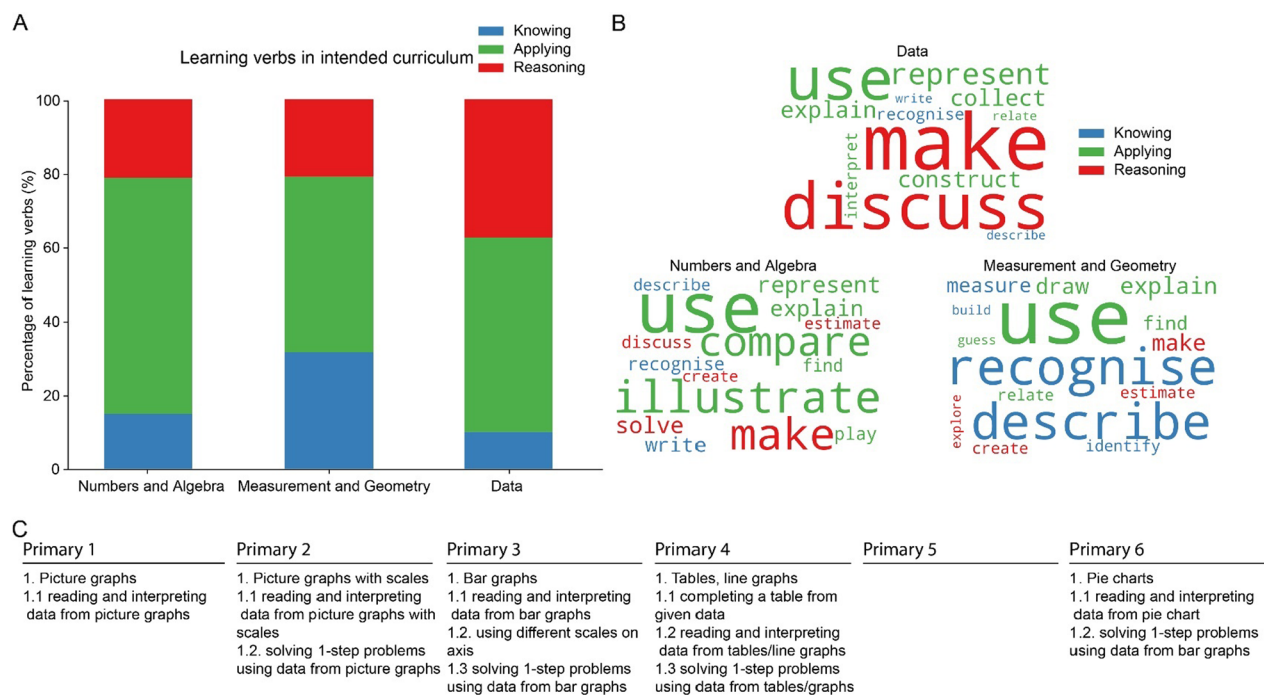


Fig. 2 Comparing mathematics content domains in the intended curriculum in Singapore. **A** Cognitive domains of learning verbs used in different content domains in the intended mathematics curriculum combined across all levels. **B** Word cloud of learning verbs as a function of content domains with size related to the frequency of word occurrence within a content domain and colors corresponding to cognitive domains. **C** Progression of contents in the data domain

Results

Intended curriculum

We first examined learning verbs used in the intended mathematics curriculum in Singapore (RO1). A total of 719 occurrences of 64 unique learning verbs were extracted and analyzed when combined across all levels. We observed that higher-order verbs related to reasoning were dominant in the data domain compared to the other content domains (Fig. 2A; RQ1a). Higher-order learning experiences were also intended for in the other two content domains but with more emphasis on lower-order verbs related to applying and knowing compared to the data domain (RQ1b). To further examine the nature of the learning experiences, we examined the frequency of learning verbs via a word cloud (Fig. 2B). In the data domain, students were, for instance, expected to “Discuss examples of data presented in various forms” and “Use the presented data display to make interpretations and predictions”. The *discuss* and *make* verbs belonging to the reasoning cognitive domain were the most frequent verbs in the data domain. In addition to the learning experiences, Fig. 2C shows the progression of data contents in the intended curriculum. The sequence is generally picture graphs (Primary 1 and 2) followed by bar graphs (Primary 3), then tables and line graphs (Primary 4). Students at Primary 6 (last elementary grade) were

introduced to pie charts (RQ1a). Overall, the results suggest that Singapore’s intended curriculum in data domain strongly emphasized learning verbs for higher cognitive skills compared to other mathematics content domains (RQ1b).

Assessed curriculum

To address RO2, we characterized 1315 summative assessment items from six public schools in Singapore. Taking items from all grade levels in total, we found that a large proportion of assessment items were devoted to higher-level reasoning skills in the data domain (Fig. 3A). We further broke down the assessment items by primary levels (grades). There were only 5.5% of items devoted to the data domain at Primary 2 and this proportion increased to 13.4% at Primary 6 (Fig. 3B; RQ2a). The numbers and algebra domain formed most items at lower primary levels while measurement and geometry became important at upper primary levels (Fig. 3B). In terms of cognitive expectations, at Primary 4, the data domain had highest emphasis on applying (Fig. 3C; RQ2b). However, from Primary 5 onward, there was a big increase in emphasis on reasoning in data (majority of all assessed items) unlike in the other two content domains where the increases in cognitive skills were more gradual as the students progressed up the primary levels (Fig. 3C;

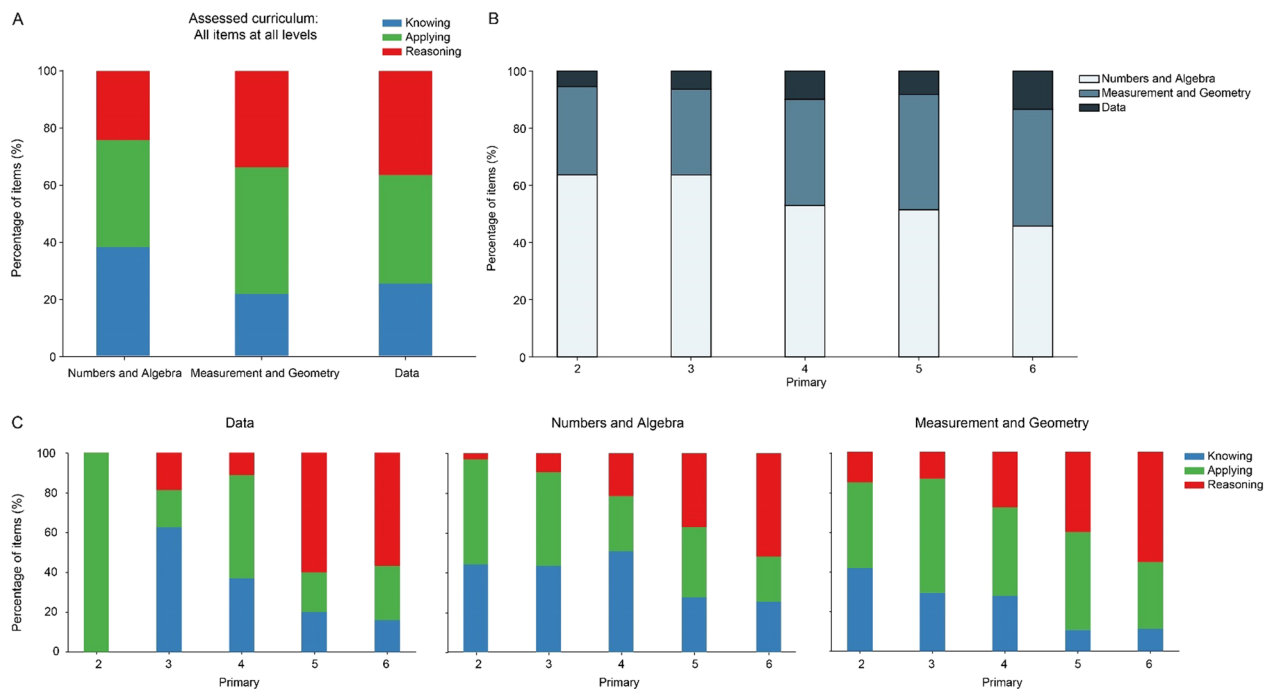


Fig. 3 Comparing mathematics content domains in the assessed curriculum in Singapore. **A** Cognitive domains in different content domains for assessment items combined across all primary levels. **B** Percentage of assessment items in different content domains from Primary 2–6. **C** Cognitive domains of assessment items from Primary 2–6 in different content domains from Primary 2–6

RQ2b). A very small proportion of items (6.4%) covered multiple content domains (RQ2c). Overall, our analysis of the assessed curriculum agrees with observations of the intended curriculum of greater proportion of higher-order cognitive skills required in the data domain compared to other mathematics content domains.

To further probe the data domain in the assessed curriculum, we quantified the types of data displays assessed (Fig. 4A). We found good alignment between the assessed items (Fig. 4A) and the intended curriculum (Fig. 2C) in terms of the progression of data visualizations at the different primary levels. For example, bar graphs were never assessed before Primary 3, aligned with the intended curriculum. There was also significant coverage at upper primary levels of data displays previously introduced at lower primary levels. For instance, even at Primary 6, there was still a good proportion of items devoted to bar graphs introduced 3 years earlier (Fig. 4A). This indicates a spiral nature of the assessed curriculum that significantly revisits previous years' contents as emphasized in the intended curriculum (Ministry of Education Singapore, 2012). There were also mixed data visualization assessment items, starting at Primary 4 (Fig. 4A). This assessment was similarly in line with the intended curriculum that emphasizes learning experiences for linking different types of data visualizations starting at Primary 4 to enhance the representational fluency of students.

Figure 4B shows examples of actual data visualizations assessed, generally covering all the contents of the intended curriculum. However, we also observed some misalignment. Tables were assessed at Primary 3 even though the official curriculum intended for it to be covered starting at Primary 4. Moreover, comparing cognitive domains in the intended data curriculum (Fig. 2A) to assessed curriculum (Fig. 3A) identified apparent misalignment such as stronger emphasis on knowing in assessed curriculum compared to intended curriculum. Overall, these observations suggested patterns of alignment with some deviations between intended and assessed curriculum.

Attained curriculum: a cross-country analysis of TIMSS achievement data

Using international large-scale assessment data from TIMSS, we addressed RO3 on how Singaporean students ($N=5041$ –6668 students from 5 cycles) performed in comparison to peers from developed East Asian countries (Hong Kong SAR, $N=2968$ –4608 students; Rep. of Korea, $N=3893$ –4334 students; Chinese Taipei, $N=3765$ –4661 students; Japan, $N=4196$ –4535 students). Figure 5 shows the distribution of scale scores. Lower quartile, median and upper quartile define each box while lower and upper whiskers represent 10th and 90th percentile, respectively. Here, we have highlighted the main patterns while Table 1

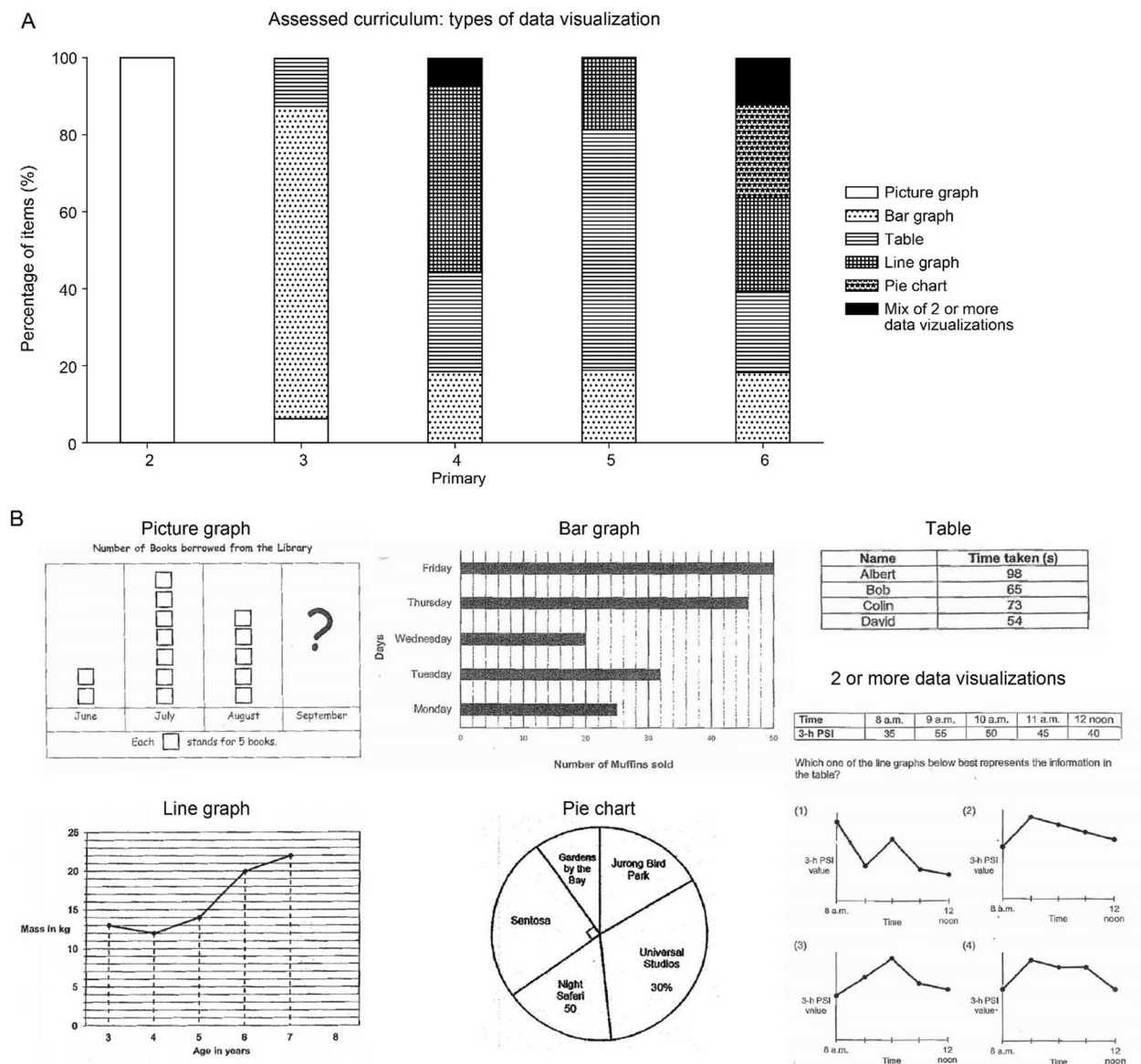


Fig. 4 Data visualizations in the assessed curriculum in Singapore. **A** Types of data visualizations assessed from Primary 2–6. **B** Examples of types of data visualizations assessed

to Table 3 have more detailed statistics of all our planned comparisons. For RQ3a, Singaporean students in the current 2019 cycle were compared to Singaporean students from previous cycles. 50th percentile Singaporean students in the current 2019 cycle performed statistically significantly better than 50th percentile Singaporean students in all four previous cycles in all three content domains (all corrected p -values < 0.001). The range of effect sizes for 2019 vs. four previous cycles for the data domain was small to medium in sizes at $d = 0.17$ – 0.60 . However, the better performance among 10th percentile Singapore students in the 2019 cycle compared to

previous cycles was less prominent than for 50th percentile Singaporean students, with small effect sizes at $d = 0.05$ – 0.11 for the data domain. When averaging the range of effect sizes, 50th percentile Singaporean students had 4.52 times 2019 performance advantage compared to 10th percentile Singaporean students in the data domain, a discrepancy that was not as large for the numbers (2.62 times), and measurement and geometry domains (2.09 times). To reiterate, the 2019 Singaporean students underwent the 2013 intended curriculum as analyzed above. In sum, analyses for RQ3a suggested a data curriculum that preferentially elevated the performance

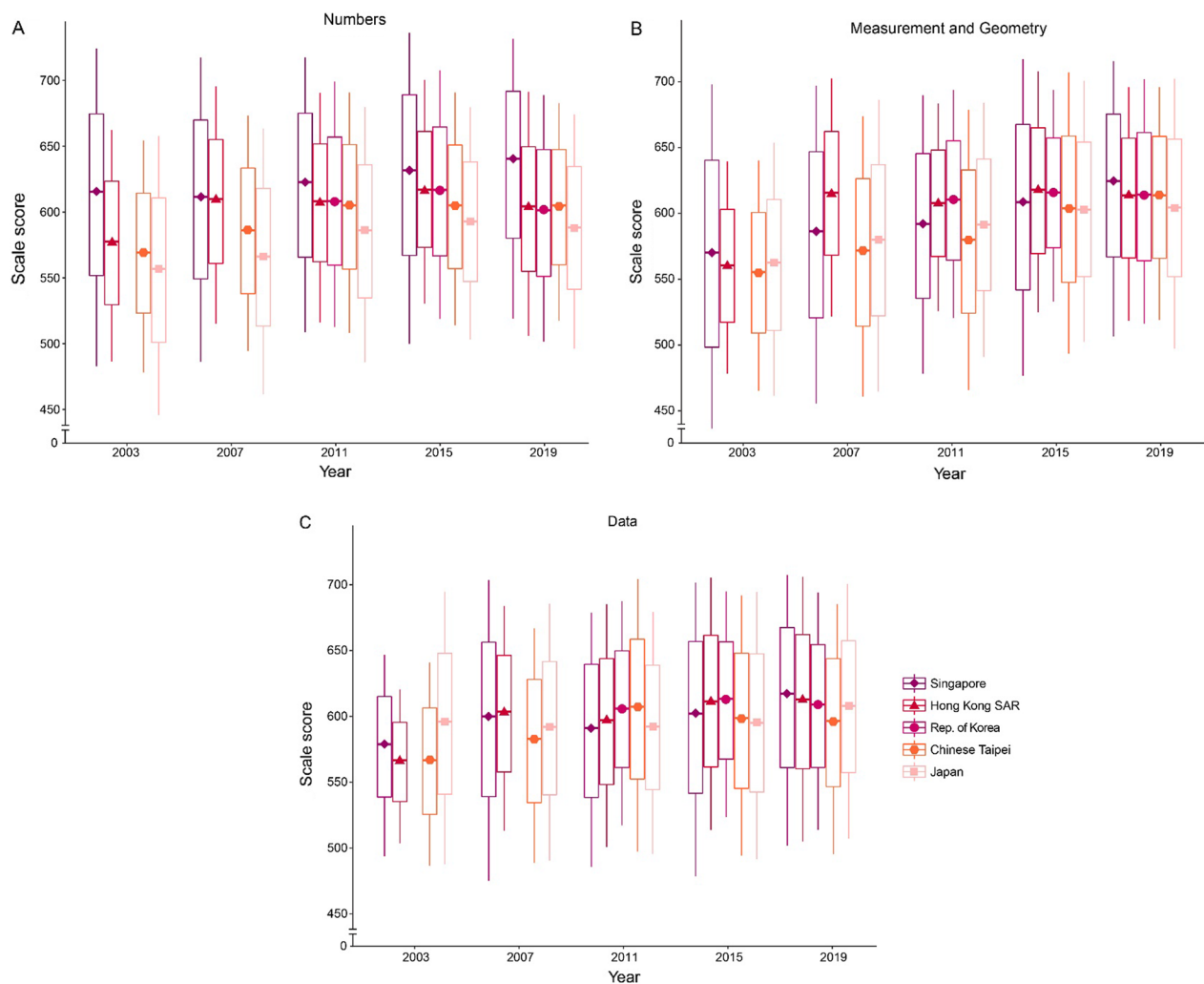


Fig. 5 Mathematics attainments of Singapore and East Asian countries in TIMSS 2003–2019. Grade 4 (equivalent to Primary 4) scale scores for **A** numbers, **B** measurement and geometry and **C** data domains. Lower quartile, median and upper quartile define each box. Lower and upper whiskers represent 10th and 90th percentile, respectively. Rep. of Korea did not participate in 2003 and 2007

of middle-performing students but had much less of a positive effect on lower-performing students. This unequal positive effect on different groups of Singaporean students was much more pronounced for the data domain than the other content domains.

To investigate the extent to which the current data curriculum was associated with increased attainments, we further examined cycle-on-cycle changes, i.e., compared change from 2015 to 2019 with 2003 to 2007, 2007 to 2011, and 2011 to 2015 (unlike the above which compared actual attainment in 2019 cycle to all other cycles). All cycle-on-cycle changes were statistically significant (p -values < 0.001), thus, we focused on effect sizes. For the 50th percentile Singaporean students, the 2015–2019 cycle-on-cycle change effect size

was $d = 0.17$ comparable to other effect sizes of 0.38 and 0.18 for 2003–2007, and 2011–2015, respectively (there was decline in performance for 2007–2011). In contrast, for the 10th percentile Singapore students, the 2015–2019 cycle-on-cycle change effect size ($d = 0.08$) was smaller than the other cycle-on-cycle changes (0.23 and 0.11 for 2007–2011 and 2011–2015, respectively; there was decline in performance for 2003–2007)). Summarizing RQ3a results, 50th percentile Singaporean students had comparable increases in 2019 attainments in the data domain across cycles despite the already high performance in 2015, likely limiting larger increases in 2019. In contrast, the curriculum produced a much smaller positive effect on the 10th percentile Singaporean students.

Table 1 Related to RQ3a, comparing attainments of 2019 Singaporean students to Singaporean students in other cycles

	Numbers		Measurement and geometry		Data	
	<i>p</i> -value	Cohen's <i>d</i>	<i>p</i> -value	Cohen's <i>d</i>	<i>p</i> -value	Cohen's <i>d</i>
50th percentile						
2003	< 0.001	− 0.27	< 0.001	− 0.55	< 0.001	− 0.60
2007	< 0.001	− 0.30	< 0.001	− 0.37	< 0.001	− 0.25
2011	< 0.001	− 0.21	< 0.001	− 0.32	< 0.001	− 0.40
2015	< 0.001	− 0.07	< 0.001	− 0.18	< 0.001	− 0.17
2019	Reference					
10th percentile						
2003	< 0.001	− 0.13	< 0.001	− 0.26	< 0.001	− 0.05
2007	< 0.001	− 0.10	< 0.001	− 0.18	< 0.001	− 0.11
2011	< 0.001	− 0.04	< 0.001	− 0.11	< 0.001	− 0.08
2015	< 0.001	− 0.05	< 0.001	− 0.09	< 0.001	− 0.08
2019	Reference					

Comparisons are made to 2019 as reference. Negative *d* represents lower performance compared to the 2019 cycle

For RQ3b, Singaporean students in the current 2019 cycle were compared to East Asian students in the same 2019 cycle. 50th percentile Singaporean students statistically significantly outperformed East Asian students in the current 2019 cycle in the data domain (all corrected *p*-values < 0.001) with medium effect sizes (*d* = 0.18 to 0.70). Similar results were obtained for other content domains comparing Singaporean to East Asian students (all *p*-values < 0.001, with medium to large effect sizes; Table 3). However, the picture for 10th percentile Singaporean students was less positive for the data domain. 10th percentile Singaporean students only statistically outperformed peers from Hong Kong and Chinese Taipei at the same percentile in the data domain (corrected *p*-values < 0.001) with small effect sizes (*d* = 0.03–0.13). 10th percentile Singaporean students did not statistically outperform students from Rep. of Korea and Japan in the same 10th percentile. This result for the data domain contrasted with a more positive picture for numbers domain in which 10th percentile Singaporean students generally outperformed East Asian peers in the same percentile (all corrected *p*-values < 0.05 except for one, with small to medium effect sizes). See Tables 1, 2, 3 for detailed statistics.

We further examined the percentage of TIMSS assessment items that matched the national curriculum. The numbers, and measurement and geometry domains had higher percentage of overlap with respective national curricula compared to the data domain (Fig. 6). Importantly, Singapore did not stand out as having a much larger test–curriculum overlap than East Asian countries in all of the content domains, including in data (Fig. 6; RQ3c). Taken together with detailed statistical analyses for RQ3a and

RQ3b, overall, the average Singaporean students generally attained very high levels of achievement relative to past cycles and relative to their East Asian peers in all content domains including in data. However, weaker Singaporean students underperformed particularly in the data domain despite the rigorous intended national data curriculum.

Discussion

Acquiring data skills has become increasingly important for the 21st century. Given that early formal learning about data is taught within mathematics, we used a curriculum approach to examine the intended, assessed,

Table 2 Related to RQ3a, comparing cycle-on-cycle changes for the data domain of Singaporean students

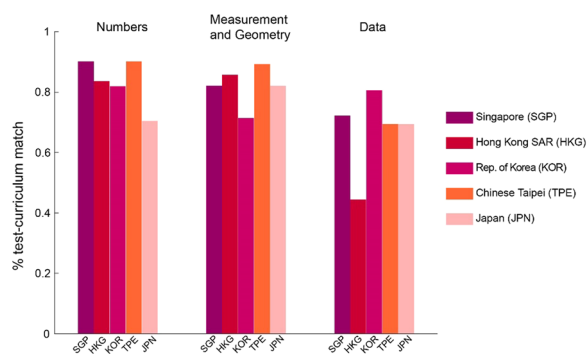
	Data	
	<i>p</i> -value	Cohen's <i>d</i>
50 th percentile		
2003–2007	< 0.001	− 0.38
2007–2011	< 0.001	0.18
2011–2015	< 0.001	− 0.20
2015–2019	< 0.001	− 0.17
10 th percentile		
2003–2007	< 0.001	0.10
2007–2011	< 0.001	− 0.23
2011–2015	< 0.001	− 0.11
2015–2019	< 0.001	− 0.08

The 2nd year is the reference cycle, e.g., 2003–2007 compares 2003 to 2007 with 2007 as a reference. Negative *d* represents an improvement in performance from the 1st year to the 2nd year

Table 3 Related to RQ3b, comparing attainments of Singaporean students in 2019 to East Asian students in 2019

	Numbers		Measurement and Geometry		Data	
	<i>p</i> -value	Cohen's <i>d</i>	<i>p</i> -value	Cohen's <i>d</i>	<i>p</i> -value	Cohen's <i>d</i>
50 th percentile						
Hong Kong	< 0.001	− 0.86	< 0.001	− 0.31	< 0.001	− 0.18
Rep. Korea	< 0.001	− 1.46	< 0.001	− 0.49	< 0.001	− 0.42
Chinese Taipei	< 0.001	− 1.16	< 0.001	− 0.42	< 0.001	− 0.70
Japan	< 0.001	− 1.83	< 0.001	− 0.73	< 0.001	− 0.39
Singapore	Reference					
10 th percentile						
Hong Kong	< 0.001	− 0.16	0.04	0.02	< 0.001	− 0.03
Rep. Korea	< 0.001	− 0.29	1	< 0.01	0.26	0.02
Chinese Taipei	< 0.001	− 0.10	0.003	0.03	< 0.001	− 0.13
Japan	< 0.001	− 0.31	< 0.001	− 0.15	1	< 0.01
Singapore	Reference					

Comparisons are made to Singapore as a reference. Negative *d* represents lower performance of comparison country to Singapore

**Fig. 6** Match between TIMSS 2019 assessment items and national curriculum

and attained mathematics curriculum. This approach is further motivated by the increasing emphasis of curriculum-related research in science, technology, engineering, and mathematics (STEM) education (Li et al., 2020). The Singapore elementary school system was used as a case study for intended and assessed curricula with detailed multi-country statistical comparisons for attained curriculum. Related to RO1, results indicated that traditional mathematics content domains strongly dominated the elementary mathematics content coverage in Singapore (RQ1a). It is also anecdotally known that Singapore teachers spend less classroom time on data compared to other content domains deemed to be more important such as numbers and measurement. Interestingly, despite the limited content coverage of the data domain, the cognitive skills assessed for data domain were high (RQ2a). Overall, both intended and

assessed data curricula placed greater emphasis on higher-order skills such as applying and reasoning compared to the other mathematics content domains, especially at higher primary levels (RQ1b, RQ2b).

A number of observations further support our claim of more demanding data curriculum compared to other content domains at Primary 4 specifically, when students were tested for actual attainment in TIMSS. First, at Primary 4, higher cognitive domains of reasoning and applying contributed a greater combined percentage of assessed items in data domain compared to numbers and algebra, though similar in proportion to measurement and geometry (Fig. 3C). Second, just as importantly, the proportion of assessed items in data domain was very small, 3–4 times smaller compared to other content domains at Primary 4. Based on previous large-scale classroom studies in Singapore, assessments in Singapore generally constrain enacted curriculum in the classroom in terms of what and how content is taught, and the opportunities to learn and practice (Hogan et al., 2013). Thus, our summative assessments data provided a window into the very limited instructional emphases and opportunities to learn about data. Third, about 10% of the assessed data items in schools at Primary 4 were mixed visualizations (Fig. 4A), generally quite challenging. Related to this point, TIMSS international assessment indeed classifies mixed data visualizations as a reasoning cognitive domain, which we also did. Overall, our observations indicated fewer opportunities to learn, practice and be assessed in data domain compared to other domains yet what were assessed consisted of quite high cognitive expectations. This rigorous data

curriculum seemed to translate to very good attainments of the average Primary 4 Singaporean student in international assessments compared to East Asian countries in the 2019 cycle (RQ3a). Prior to 2019, the average Singaporean students did not top East Asian countries in the data domain, suggesting a demanding curriculum that might be more recent (RQ3b). Thus, Singapore provides an interesting case study of how learning intentions and assessments can remain challenging despite much smaller content coverage of data domain.

Previous work has examined how students understand data and its visualization, with a focus on learning about graphs, which tend to be the most common data visualization in K-16 education (Aksoy & Bostan, 2021; Friel et al., 2001). A graph is made up of many different types of symbols: geometric such as lines and points, linguistic such as words and numerals, and pictorial such as icons (Börner et al., 2019). These symbols are spread across a relatively large spatial layout, all of which must be integrated cognitively for the task at hand. According to Carpenter and Shah (1998), graph comprehension is made up of 3 stages: pattern recognition stage to chunk information (e.g., *x-axis* vs. *y-axis*), a stage involving interpretation of the relationship in graph data (e.g., trends in a line graph) and another interpretative stage involving referents (e.g., axes labels). These stages are cyclically and incrementally integrated instead of being a strictly serial process, with more complex graphs requiring more time for cycles of integration. More complex reasoning skills such as predicting the next data point in the graph likely also involve more cycles of integration for a coherent understanding. In summary, developing data literacy skills can be challenging as it places higher levels of cognitive skills on the reader. When coupled with the lack of content emphasis in formal mathematics curriculum, it is not surprising then that even educated adults can exhibit difficulties in reading and understanding data (Börner et al., 2016; Kaplar et al., 2021).

Our results suggest that lower-achieving Primary 4 Singaporean students underperform in the data domain. While the average Singaporean student outperformed the average East Asian student at Primary 4, the TIMSS data also suggested that weaker Singaporean students were not benefiting from the curriculum as much, thus not ranking as favorably compared to weaker East Asian students (RQ3a). One hypothesis is that the demanding data curriculum, while significantly enhancing the performance of the average Singaporean student, is not able to meet the learning needs of the weaker students who fall farther behind. The skills demanded of data literacy as elaborated above may not be sufficiently developed in weaker students. Our hypothesis is supported by a recent study that used item response theory, finding that

data-related items in a science inference instrument were particularly difficult for lower-track Singaporean students compared to other items (Teo & Goh, 2019). Our results are also relevant to recent debates over equity–excellence trade-off which posits that higher average performance necessitates a more unequal educational system (Van de Werfhorst & Mijs, 2010). The extent to which this trade-off is empirically supported has been questioned (Parker et al., 2018). Yet, at least in the case of Singapore, the high average performance does seem to come at the price of a much wider distribution of scores such that the tail end of the distribution for Singaporean students is lower than for East Asian countries. Our study provides motivation to further compare the learning experiences involving average and weaker students in the context of a seemingly rigorous data curriculum. A hypothesis based on the above model of cognitive integration is that weaker students, while able to identify the disparate symbols and elements of data visualization, fail to engage in repeated cognitive integrative cycles required to form a fuller understanding in order to solve the problem at hand. Further studies of the learning processes, especially among lower-performing students, would be useful.

Our comparative cross-content domain approach is also pertinent to current efforts in data science and statistics education. We observed two seeming trends in this area. First, many have argued that statistical thinking is quite distinct from mathematics. Real-world contexts, variation in data, chance, and uncertainty all play prominent roles in statistical meaning-making but these are abstracted away in mathematical thinking as they obscure pure mathematical structures (Cobb & Moore, 1997). This approach has led to success in growing the field of data science and statistics education as a field of inquiry worthy of its own standing (Ben-Zvi et al., 2017; Leavy et al., 2018). The other trend we observed is that STEM frameworks have incorporated data and visualizations as key pillars for an integrated STEM learning experience (Kelley & Knowles, 2016). While these trends are welcomed as they make data science more prominent in K-12 education in a manner that cuts across traditional academic subjects, the reality is that students first formally learn about data via the subject of mathematics in many countries (Davies & Sheldon, 2021; Groth, 2018). This situation has major implications. It can make learning about data subservient to the broader mathematics curriculum dominated by other content domains as found in this study. Moreover, students may be influenced to think about data in ways that highlight mathematical focus on algorithms, abstraction and problem-solving with deterministic answers (Bargagliotti & Groth, 2016). Further, Davies and Sheldon (2021) shared an anecdote of a mathematics assessment meeting in which an item

about normal distribution had a grading scheme that rewarded both “yes” and “no” answers as long as correct justifications were given. Mathematics teachers protested such uncertainty, which is not typically tolerated in traditional mathematics content domains (p. S59). Our study is limited by the nature of the curriculum documents and data that were not explicitly aimed at addressing clashes between statistical and mathematical thinking. Nonetheless, we have identified differences in learning expectations and assessment between data and other mathematical domains that can be useful when further comparing different educational systems.

While many promote data science as cutting *across* disciplines, a complementary effort is to examine how to apply skills associated with data *within* mathematics itself. Cross-content domain learning was found to be very limited in this study as very few items assessed multiple content domains (6.4%), even fewer involving the data domain. This finding is consistent with a previous study on textbook curriculum (Jones et al., 2015). Thus, students would potentially have missed out on important opportunities to learn the useful cross-applicability of data, a key aspect of data science, suggesting the need to further integrate data skills and competencies into the mathematics curriculum. There are different strategies to do so. For example, in measurement and geometry, elementary school children are taught how to measure areas deterministically via formulas. However, Monte Carlo methods using random numbers exist to estimate areas, particularly useful on odd shapes. A older study emphasized how students can discover value of π using Monte Carlo methods (Easterday & Smith, 1991) and now such learning experiences can be easily incorporated using modern statistical software (e.g., Fitzallen & Watson, 2010). Other examples include using data presented in various forms to enhance student’s representational fluency in mathematics content domains such as functions (Ceuppens et al., 2018). Students can be taught, within mathematics, how to think in a data-driven manner, graphically visualize the data, and make links to traditional mathematical solutions (often via formulas). We consider such cross-content domain learning to be analogous to near transfer within mathematics in contrast to far transfer when applying data science across traditional disciplines (Roehrig et al., 2021). Our proposed call for near transfer efforts agree with previous views that emphasize the enriching role of data and statistics in mathematics (Davies et al., 2012; Goldstein, 2007).

The present study has limitations. There is always a risk in taking a comparative approach. First, we developed a framework for comparing the intended and

assessed curriculum across very different mathematics content domains. One might argue it is difficult to compare cognitive skills for topics as distinct as, for example, bar graphs in data and angles in geometry. Nonetheless, we believe that our approach, based upon previous work on the thinking curriculum, further used by the TIMSS assessment framework, is general enough to be applicable across content domains. Second, in taking multi-country comparisons, there are issues in comparing attainments across quite different educational systems and contexts. Curriculum coverage is one issue. Here, it did not appear that test–curriculum match played a major role in determining broad patterns of national attainments, a result consistent with more detailed TIMSS research (Mullis et al., 2020). However, we cannot rule out more subtle relationships between test–curriculum overlap and attainments as scores of individual test items could not be analyzed. Moreover, our study was limited to intended, assessed, and attained curriculum. Future studies can examine enacted curriculum as well as hidden aspects of curriculum involving attitudes and values when learning about data within the context of mathematics.

Conclusion

One of the emerging forms of literacies is data literacy, important in an increasingly data-driven world. Students first formally learn about data within the elementary mathematics curriculum but there is a gap in knowledge on how they learn so in relation to other foundational mathematics content domains. Using a curriculum framework, we analyzed the intended, assessed, and attained curricula in the data domain compared to other mathematics content domains such as numbers, algebra, measurement, and geometry using Singapore as a case study. Findings suggested that, despite very limited coverage, the data domain required a greater proportion of higher-order cognitive skills than other content domains in both intended and assessed curricula. Moreover, this data curriculum was associated with high performance by the average Singaporean student compared to East Asian students using international large-scale assessment data. However, lower-achieving Singaporean students lagged behind their East Asian peers, especially in the data domain, with implications on current equity issues in STEM education. Moreover, the very limited cross-domain applications of data highlight the need for elementary school students to be exposed to learning experiences that emphasize the cross-applicability of data, especially in the mathematics curriculum.

Abbreviations

GAISE II	Guidelines for Assessment and Instruction in Statistics Education II
SA2	Semestral assessment 2
STEM	Science, technology, engineering, and mathematics
TCMA	Test–Curriculum Matching Analysis
TIMSS	Trends in International Mathematics and Science Study

Acknowledgements

TIMSS dataset courtesy of IEA's Trends in International Mathematics and Science Study (TIMSS) Copyright © 2021 International Association for the Evaluation of Educational Achievement (IEA).

Author contributions

FA and YKOW conceived, designed, and implemented the research. FA and YKOW wrote the manuscript. IHY contributed to interpretation of the data, and writing and revision of manuscript. All authors read and approved the final manuscript.

Funding

We wish to acknowledge the funding support for this project from Nanyang Technological University under the URECA program.

Availability of data and materials

The datasets generated and/or analyzed during the current study are available in the NIE Data Repository, <https://doi.org/10.25340/R4/YVQMVF>

Declarations

Competing interests

The authors declare that they have no competing interests.

Received: 13 June 2022 Accepted: 6 January 2023

Published online: 30 January 2023

References

- Aksoy, E. Ç., & Bostan, M. I. (2021). Seventh graders' statistical literacy: An investigation on bar and line graphs. *International Journal of Science and Mathematics Education*, 19(2), 397–418.
- Ali, F. (2016). Gaps in educational outcomes: Analysing national examination performance of Singaporean Malay and non-Malay students in the past 20 years. *Asia Pacific Journal of Education*, 36(4), 473–487.
- Alsubaie, M. A. (2015). Hidden curriculum as one of current issue of curriculum. *Journal of Education and Practice*, 6(33), 125–128.
- Ang, R. P., Li, X., Huan, V. S., Liem, G. A. D., Kang, T., Wong, Q., & Yeo, J. Y. (2020). Profiles of antisocial behavior in school-based and at-risk adolescents in Singapore: A latent class analysis. *Child Psychiatry and Human Development*, 51(4), 585–596.
- Bakker, A., Cai, J., & Zenger, L. (2021). Future themes of mathematics education research: An international survey before and during the pandemic. *Educational Studies in Mathematics*, 107(1), 1–24.
- Bargagliotti, A., Franklin, C., Arnold, P., Gould, R., Johnson, S., Perez, L., & Spangler, D. (2020). *Pre-K-12 guidelines for assessment and instruction in statistics education (GAISE) report II*. American Statistical Association and National Council of Teachers of Mathematics.
- Bargagliotti, A., Arnold, P., & Franklin, C. (2021). GAISE II: Bringing data into classrooms. *Mathematics Teacher Learning and Teaching*, 114(6), 424–435.
- Bargagliotti, A., & Groth, R. (2016). When mathematics and statistics collide in assessment tasks. *Teaching Statistics*, 38(2), 50–55.
- Batanero, C., Burrill, G., & Reading, C. (2011). *Teaching statistics in school mathematics-challenges for teaching and teacher education: A joint ICMI/ IASE study: The 18th ICMI study*. Springer.
- Ben-Zvi, D., Makar, K., & Garfield, J. (2017). *International handbook of research in statistics education*. Springer.
- Biehler, R., Frischemeier, D., Reading, C., & Shaughnessy, J. M. (2018). Reasoning about data. In D. Ben-Zvi, K. Makar, & J. Garfield (Eds.), *International handbook of research in statistics education* (pp. 139–192). Springer.
- Blei, D. M., & Smyth, P. (2017). Science and data science. *Proceedings of the National Academy of Sciences of the United States of America*, 114(33), 8689–8692.
- Börner, K., Bueckle, A., & Ginda, M. (2019). Data visualization literacy: Definitions, conceptual frameworks, exercises, and assessments. *Proceedings of the National Academy of Sciences of the United States of America*, 116(6), 1857–1864.
- Börner, K., Maltese, A., Balliet, R. N., & Heimlich, J. (2016). Investigating aspects of data visualization literacy using 20 information visualizations and 273 science museum visitors. *Information Visualization*, 15(3), 198–213.
- Bråting, K., & Kilhamn, C. (2021). The integration of programming in Swedish school mathematics: Investigating elementary mathematics textbooks. *Scandinavian Journal of Educational Research*, 78, 1–16.
- Cao, L. (2017). Data science: A comprehensive overview. *ACM Computing Surveys (CSUR)*, 50(3), 1–42.
- Caro, D. H., & Biecek, P. (2017). intsvy: An R package for analyzing international large-scale assessment data. *Journal of Statistical Software*, 81(1), 1–44.
- Carpenter, P. A., & Shah, P. (1998). A model of the perceptual and conceptual processes in graph comprehension. *Journal of Experimental Psychology: Applied*, 4(2), 75–100.
- Ceuppens, S., Deprez, J., Dehaene, W., & De Cock, M. (2018). Design and validation of a test for representational fluency of 9th grade students in physics and mathematics: The case of linear functions. *Physical Review Physics Education Research*, 14(2), 020105.
- Chen, Q. (2014). Using TIMSS 2007 data to build mathematics achievement model of fourth graders in Hong Kong and Singapore. *International Journal of Science and Mathematics Education*, 12(6), 1519–1545.
- Chen, W.-L., Elchert, D., & Asikin-Garmager, A. (2018). Comparing the effects of teacher collaboration on student performance in Taiwan. *Hong Kong and Singapore. Compare*, 50(4), 515–532.
- Chia, H. T. (2016). Students' sense-making of graphical representation in a basic statistics module. In D. Ben-Zvi & K. Makar (Eds.), *The Teaching and Learning of Statistics* (pp. 177–178). Springer.
- Cobb, G. W., & Moore, D. S. (1997). Mathematics, statistics, and teaching. *American Mathematical Monthly*, 104(9), 801–823.
- Curcio, F. R. (1987). Comprehension of mathematical relationships expressed in graphs. *Journal for Research in Mathematics Education*, 18(5), 382–393.
- Davies, N., Marriott, J. M., & Bidgood, R. G. P. (2012). *Teaching statistics in British secondary schools: Statistics knowledge and pedagogy in secondary mathematics teacher training courses in British higher education institution*. T. S. Trust.
- Davies, N., & Sheldon, N. (2021). Teaching statistics and data science in England's schools. *Teaching Statistics*, 43, 552–570.
- Dingman, S., Teuscher, D., Newton, J. A., & Kasmer, L. (2013). Common mathematics standards in the United States: A comparison of K–8 state and Common Core standards. *The Elementary School Journal*, 113(4), 541–564.
- Donoho, D. (2017). 50 years of data science. *Journal of Computational and Graphical Statistics*, 26(4), 745–766.
- Easterday, K., & Smith, T. (1991). A Monte Carlo application to approximate pi. *The Mathematics Teacher*, 84(5), 387–390.
- Engel, J. (2017). Statistical literacy for active citizenship: A call for data science education. *Statistics Education Research Journal*, 16(1), 44–49.
- Fishbein, B., Foy, P., & Yin, L. (2021). *TIMSS 2019 User Guide for the International Database* (2nd ed.). Boston College, TIMSS & PIRLS International Study Center.
- Fitzallen, N., & Watson, J. (2010). Developing statistical reasoning facilitated by TinkerPlots. In C. Reading (Ed.), *Data and context in statistics education: Towards an evidence-based society. Proceedings of the Eighth International Conference on Teaching Statistics (ICOTS8)*.
- Friel, S. N., Curcio, F. R., & Bright, G. W. (2001). Making sense of graphs: Critical factors influencing comprehension and instructional implications. *Journal for Research in Mathematics Education*, 32(2), 124–158.
- Geisinger, K. F. (2016). 21st century skills: What are they and how do we assess them? *Applied Measurement in Education*, 29(4), 245–249.
- Goldstein, H. (2007). The future of statistics within the curriculum. *Teaching Statistics: An International Journal for Teachers*, 29(1), 8–9.
- Groth, R. E. (2018). Unpacking implicit disagreements among early childhood standards for statistics and probability. In A. Leavy, M. Meletiou-Mavrotheris, & E. Paparistodemou (Eds.), *Statistics in early childhood and primary education* (pp. 149–162). Springer.

- Hirsch, C. R., & Reys, B. J. (2009). Mathematics curriculum: A vehicle for school improvement. *ZDM Mathematics Education*, 41(6), 749–761.
- Hogan, D., Chan, M., Rahim, R., Kwek, D., Maung Aye, K., Loo, S. C., Sheng, Y. Z., & Luo, W. (2013). Assessment and the logic of instructional practice in Secondary 3 English and mathematics classrooms in Singapore. *Review of Education*, 1(1), 57–106.
- Johansson, S. (2020). Analysing the (mis)use and consequences of international large-scale assessments. In J. Zajda (Ed.), *Globalisation, Ideology and Education Reforms* (Globalisation, Comparative Education and Policy Research, vol 20, pp. 13–24). Springer, New York.
- Jones, D. L., Brown, M., Dunkle, A., Hixon, L., Yoder, N., & Silbernack, Z. (2015). The statistical content of elementary school mathematics textbooks. *Journal of Statistics Education*, 23(3), 89.
- Kaplar, M., Lužanin, Z., & Verbić, S. (2021). Evidence of probability misconception in engineering students—why even an inaccurate explanation is better than no explanation. *International Journal of STEM Education*, 8(1), 1–15.
- Karpatne, A., Atluri, G., Faghmous, J. H., Steinbach, M., Banerjee, A., Ganguly, A., Shekhar, S., Samatova, N., & Kumar, V. (2017). Theory-guided data science: A new paradigm for scientific discovery from data. *IEEE Transactions on Knowledge and Data Engineering*, 29(10), 2318–2331.
- Kelley, T. R., & Knowles, J. G. (2016). A conceptual framework for integrated STEM education. *International Journal of STEM Education*, 3(1), 1–11.
- Kelly, A. V. (2009). *The curriculum: Theory and practice*. Sage.
- Krathwohl, D. R. (2002). A revision of Bloom's taxonomy: An overview. *Theory into Practice*, 41(4), 212–218.
- Leavy, A., Meletiou-Mavrotheris, M., & Paparistodemou, E. (2018). *Statistics in early childhood and primary education*. Springer.
- Lee, N. H., Ng, W. L., & Lim, L. G. P. (2019). The intended school mathematics curriculum. In T. L. Toh & B. Kaur (Eds.), *Mathematics education in Singapore* (pp. 35–53). Springer: Berlin.
- Li, Y., Wang, K., Xiao, Y., & Froyd, J. E. (2020). Research and trends in STEM education: A systematic review of journal publications. *International Journal of STEM Education*, 7(1), 1–16.
- Lindquist, M., Philpot, R., Mullis, I. V. S., & Cotter, K. E. (2017). Chapter 1 - TIMSS 2019 mathematics framework. In I. V. S. Mullis & M. O. Martin (Eds.), *TIMSS 2019 assessment frameworks* (pp. 1–25). IEA Boston College, TIMSS & PIRLS International Study Center.
- Lovett, J. N., & Lee, H. S. (2017). New standards require teaching more statistics: Are preservice secondary mathematics teachers ready? *Journal of Teacher Education*, 68(3), 299–311.
- Lv, S.-H., & Cao, C. (2018). The evolution of mathematics curriculum and teaching materials in secondary schools in the twenty-first century. In Y. Cao & F. Leung (Eds.), *The 21st century mathematics education in China* (pp. 147–169). Springer.
- Ministry of Education Singapore. (2012). *Mathematics syllabus primary one to six (implementation starting with 2013 primary one cohort)*.
- Mullis, I. V. S., Martin, M. O., Foy, P., Kelly, D. L., & Fishbein, B. (2020). *TIMSS 2019 international results in mathematics and science*. TIMSS & PIRLS International Study Center.
- Nisbet, J. (2005). The thinking curriculum. In *Subject Learning in the Primary Curriculum* (pp. 286–297). Routledge.
- Parker, P. D., Marsh, H. W., Jerrim, J. P., Guo, J., & Dicke, T. (2018). Inequity and excellence in academic performance: Evidence from 27 countries. *American Educational Research Journal*, 55(4), 836–858.
- Roehrig, G. H., Dare, E. A., Ring-Whalen, E., & Wieselmann, J. R. (2021). Understanding coherence and integration in integrated STEM curriculum. *International Journal of STEM Education*, 8(1), 1–21.
- Schleicher, A. (2019). *PISA 2018: Insights and interpretations*. Berlin: OECD Publishing.
- Shin, D. (2021). Preservice mathematics teachers' selective attention and professional knowledge-based reasoning about students' statistical thinking. *International Journal of Science and Mathematics Education*, 19(5), 1037–1055.
- Stanny, C. J. (2016). Reevaluating Bloom's Taxonomy: What measurable verbs can and cannot say about student learning. *Education Sciences*, 6(4), 37.
- Tan, C. (2018). *Comparing high-performing education systems: Understanding Singapore, Shanghai, and Hong Kong*. Routledge.
- Tanudjaya, C. P., & Doorman, M. (2020). Examining higher order thinking in Indonesian lower secondary mathematics classrooms. *Journal on Mathematics Education*, 11(2), 277–300.
- Teo, T. W., & Goh, W. P. J. (2019). Assessing lower track students' learning in science inference skills in Singapore. *Asia-Pacific Science Education*, 5(1), 1–19.
- Toh, T. L., Kaur, B., & Tay, E. G. (2019). *Mathematics education in Singapore*. Springer.
- Van de Werfhorst, H. G., & Mijs, J. J. (2010). Achievement inequality and the institutional structure of educational systems: A comparative perspective. *Annual Review of Sociology*, 36, 407–428.
- Watson, J. M. (2017). Linking science and statistics: Curriculum expectations in three countries. *International Journal of Science and Mathematics Education*, 15(6), 1057–1073.
- Wild, C. J., Utts, J. M., & Horton, N. J. (2018). What is statistics? In D. Ben-Zvi, K. Makar, & J. Garfield (Eds.), *International handbook of research in statistics education* (pp. 5–36). Springer.
- Wise, A. F. (2020). Educating data scientists and data literate citizens for a new generation of data. *Journal of the Learning Sciences*, 29(1), 165–181.
- Wu, Y., & Wong, K. Y. (2007). Impact of a spreadsheet exploration on secondary school students' understanding of statistical graphs. *Journal of Computers in Mathematics and Science Teaching*, 26(4), 355–385.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)