MODESTUM
EURASIA Journal of Mathematics, Science and Technology Education, 2024, 20(4), em2421
ISSN:1305-8223 (online)
OPEN ACCESS
Research Paper
https://doi.org/10.29333/ejmste/14358

A quantitative case study of secondary school students' level of statistical thinking

Salbiah Mohamad Hasim [1], Roslinda Rosli [1*], Lilia Halim [1]

[1] The National University of Malaysia, Bangi, MALAYSIA

Received 07 August 2023 ▪ Accepted 16 December 2023

Abstract

In the 21st century, statistical thinking has become necessary for all citizens as this skill is vital for societal literacy. Students with statistical thinking will grasp and apply the problem's context to develop research and draw conclusions, as well as the coherence of the whole process from asking questions to collecting data, evaluating, and testing hypotheses. Although statistical thinking is increasingly in demand in various emerging vocations, students and teachers find statistics challenging to understand. Thus, the study examines the level of statistical thinking of the process of describing data, organizing and reducing data, representing data, and analyzing and interpreting data among high school students based on gender. The statistical thinking test modification of the framework validation test, code, and subprocess reference was used to collect research data from 35 grade 10 students. The research data were statistically analyzed using the statistical package for social sciences version 23 software through descriptive and inferential statistics, specifically an independent t-test analysis. The findings revealed that none of the students achieved an analytical level of understanding; instead, their general statistical thinking skills were at the transitional level, followed by the idiosyncratic and quantitative levels. The findings also demonstrated no statistically significant gender-related disparities in students' statistical thinking. The study proposes several recommendations, including the necessity of connecting statistical activities to the reality of students' life and area of study and emphasizing practical rather than theoretical aspects.

Keywords: data and chance, statistical thinking, secondary school students, statistical literacy, statistics education

INTRODUCTION

Global crises, such as the COVID-19 pandemic, climate change, and Industrial Revolution 4.0, have affected and changed the economic landscape, employment, and education systems (Aristovnik et al., 2020; Schleicher, 2022; United Nations, 2021; World Bank, 2022). The demand for human capital and the global job market are also experiencing significant changes. The need for the field of statistics is increasing significantly due to the emergence of several new jobs, such as artificial intelligence specialist, data scientist, data engineer, big data developer, and data analyst (World Economic Forum, 2020).

In daily life situations, people utilize statistical information to make judgements and predictions in many sectors of daily life, from consumer items to sports, weather, and forecasts. Statistics are increasingly needed due to the emergence of new job industries, such as data science, data engineering, big data development, and data analysis (World Economic Forum, 2020). As a result, statistical thinking is becoming a necessary ability for all citizens. Indeed, statistical thinking comprises more than merely performing statistical calculations or defining ideas; it also includes analyzing, reasoning, deducing, and generalizing data (Altaylar & Kazak, 2021; Langrall & Mooney, 2002; Le, 2017; Mooney, 2002).

The development of statistical thinking starts in the early grades and continues through and beyond high school, according to mathematics content standards. Students are becoming experienced in improving their statistical competency and knowledge and their abilities

✉ p109006@siswa.ukm.edu.my ✉ roslinda@ukm.edu.my **(*Correspondence)** ✉ lilia@ukm.edu.my

Contribution to the literature

- This study addresses a gap in the existing literature to the best of the researchers' knowledge and, based on a thorough literature review, investigates statistical thinking among secondary school students in Malaysia in four processes: describing data, organizing and reducing data, presenting data, and analyzing and interpreting data.
- This study will add to the body of knowledge in the field of statistical literacy by providing crucial baseline data on the statistical thinking levels of Malaysian secondary school students. The data can be used to inform future research and policy choices.
- Moreover, this study could contribute to the development of strategies to enhance statistical thinking among secondary school students by comprehending, organizing, condensing, and presenting data in various ways. Mathematics teachers can use the findings of this study to avoid gender bias when teaching statistics.

in statistics and data science. Statistical thinking requires students to possess more than a mere understanding of ideas and methods and to complete computations. Given the importance of statistical data in contemporary life, it is critical to develop statistical thinking abilities throughout school mathematics instruction.

Statistical thinking is the developmental level of students' thinking in statistics. It involves four processes: describing, organizing and classifying, representing, analyzing, and interpreting data (Langrall & Mooney, 2002; Mooney, 2002). In addition, Langrall and Mooney (2002) stated that students experience four levels of thinking within each process: idiosyncratic, transitional, quantitative, and analytical. These levels are constructed from one or more dimensions that explain the link between system components and their processes and indicate the possibilities of their routes (Langrall & Mooney, 2002). Based on these findings, the researchers believed that the statistical thinking framework represents a simplified concept that describes, clarifies, summarizes, and analyses the nature of statistical thinking, skills, and levels and the relationships between them. Additionally, it clarifies the processes of planning, implementing, and evaluating statistical thinking in sequential or separate steps to achieve outlined goals.

Problem Statement

Many students and teachers find it challenging to learn statistics (Lavidas et al., 2020). Many students need help understanding arithmetic in statistics (fractions, decimals, and algebra). They also avoid diverse interpretations based on statistical assumptions (Garfield et al., 2015; Le, 2017). Indeed, statistical questions are complex, and students are encouraged to rely more on frequent and inaccurate experiences and intuition than on statistical approaches. In addition, students have low motivation, fear, and negative perceptions and underestimate the usefulness of learning statistical concepts (Bromage et al., 2022; Levpuscek & Cukon, 2022; Rohana & Ningsih, 2019; Setambah et al., 2019; Syed Zamri et al., 2020).

The most common misconceptions in students' statistical thinking are misreading ideas, misinterpreting descriptive information, adopting incorrect techniques, and using incomplete information. Among the possible drivers of student fallacy are the assimilation of statistical ideas into an unsuitable schema, the inability to utilize knowledge sources, and the lack of capacity to integrate and incorporate information from multiple sources. Recent recommendations have also emphasized the importance of increasing students' awareness of statistical concepts, their applications, and their understanding of the multifaceted statistical concepts that have become part of various aspects of the subject matter (Nemrawi et al., 2022).

The discipline of statistics is the primary area of interest for those who work as professional statisticians. In order to have a comprehensive understanding of the subject matter, it is essential for students to actively participate in the practice. However, students were found to be unable to think statistically or master the processes that statistics practice. Among the issues raised are engaging in collaborative problem-solving endeavors that intersect the fields of mathematics and statistics. The process involves the interpretation and reinterpretation of issue settings and questions, as well as the interpretation, organization, and manipulation of data throughout the building of a model. and making casual deductions (English & Watson, 2018). Studies involving students' competence in statistical thinking processes are often carried out at the primary school level (Altaylar & Kazak, 2021; English & Watson, 2018; Kazak et al., 2018; Smith et al., 2019; Watson et al., 2020) but few focus on secondary school students.

According to these studies, some students are still struggling with comparing plots, making justifications, and identifying typical values to make predictions (Watson et al., 2020). For each process for example describing data, the majority of students are at the idiosyncratic level, whereas for the process of organizing and reducing data most students can reach the quantitative level, where the students would demonstrate awareness of relevant display features of the data (Altaylar & Kazak, 2021).

Besides, the study also found that when asked to describe the form of the plots, the girls' school demonstrates this natural inclination connected to variation first. Many more females cite variety in frequencies or data dispersion than central expectation (Watson et al., 2020). Frequently, the circumstances have been associated with the comparison of traits between females and males or other forms of interventions. These particular settings provide scenarios in which the anticipated outcomes for each group vary. As an instance, it is anticipated that the quantity of meaningful words retained in memory would surpass the quantity of nonsensical words retained or that the response time of year 10 male students would be faster compared to that of year 5 male students. Frequently, in such circumstances, students tend to prioritize examination of the mean or median disparity while giving little attention to the variability present in the two datasets (Watson, 2013).

Moreover, the study discovered a fragile positive linear relationship between the components of cognitive competence and value. Overall, it demonstrated that students' statistical reasoning ability could have been improved as, at the time of the study, they had no interest, knowledge, or skills in the use of statistics or how it can provide value in their daily lives. Students may assess statistical reasoning better with a positive attitude towards statistical knowledge, skills, relevance, and usefulness (Saidi & Siew, 2022). Rohana and Ningsih (2019) explained that most students have the first level of statistical reasoning. This level will increase when teachers use aids such as environmental technology (Conway et al., 2019; Ganesan & Kwan Eu, 2020). Thus, the guiding research questions are the following:

1. What is secondary school students' level of statistical thinking based on the four skills of describing data, organizing and reducing data, representing data, and analyzing and interpreting data?
2. To what extent does the level of statistical thinking differ between male and female students?

METHODOLOGY

Research Design

The study uses the design of a survey consisting of 35 students randomly selected from four schools in the same district. This study is a pilot study of the intervention approach of STEM integration in learning statistics involving students at secondary schools. In the context of pilot research, sample sizes ranging from 10 to 30 are deemed adequate to conduct hypothesis testing (Hill, 1998). The statistical thinking test (STT) questionnaire used is a modification of the framework validation test, code, and subprocess reference (Mooney, 2002). There are four statistical thinking processes that consist of describing data, organizing and reducing data, representing data, and analyzing and interpreting data (**Table 1**). There are six items for describing data (D1 to D6), five items for organizing and reducing data (O7 to O11), six items for representing data (R12 to R17), and eleven items for analyzing and interpreting data (A18 to A28). It comprised 28 items intended to measure statistical thinking processes and skill levels.

Table 1. CVR survey of STTs

Item	E1	E2	E3	E 4	E5	E6	E7	CVR
D1	1	1	1	1	1	1	1	1.00
D2	1	1	1	1	1	1	1	1.00
D3	1	0	1	1	1	1	1	0.75
D4	1	1	1	1	1	1	1	1.00
D5	1	1	1	1	1	1	1	1.00
D6	1	1	0	1	1	1	1	0.75
O7	0	1	1	1	1	1	1	0.75
O8	1	1	0	1	1	1	1	0.75
O9	1	0	1	1	1	1	1	0.75
O10	1	1	1	1	1	1	1	1.00
O11	1	1	1	1	1	1	0	0.75
R12	1	1	1	1	1	1	1	1.00
R13	1	1	1	1	1	1	1	1.00
R14	1	1	1	1	1	1	1	1.00
R15	1	0	1	1	1	1	1	0.75
R16	1	0	1	1	1	1	1	0.75
R17	1	1	1	1	1	1	1	1.00
A18	0	1	1	1	1	1	1	0.75
A19	1	1	1	1	1	1	1	1.00
A20	1	1	1	1	0	1	1	0.75
A21	1	1	0	1	1	1	1	0.75
A22	1	1	1	1	1	1	1	1.00
A23	1	1	1	0	1	1	1	0.75
A24	1	1	1	1	0	1	1	0.75
A25	1	1	1	1	1	1	1	1.00
A26	1	1	1	0	1	1	1	0.75
A27	1	1	1	1	0	1	1	0.75
A28	1	1	1	1	1	1	0	0.75

Several procedures were implemented to obtain validity. Language experts were approached to help ascertain the terms and language to suit the target group. Seven experts in statistics and mathematics subsequently helped to modify vague and difficult-to-understand items and terms. For content validity, however, after obtaining the approval and assessment of the expert panel, the validity of the contents of the survey instrument was measured by the procedure of the quantitative measurement of content validity by Lawshe (1975), namely the content validity ratio (CVR). CVR is used to measure the validity of item content through empirical measurement (Matore et al., 2017). Following this CVR calculation helps the examiner decide to retain or discard an item on the instrument. CVR also aims to empirically filter items on instruments with quantitative procedures to ensure that each item actually represents the contents of the construction domain (Yusoff, 2019). CVR value is in the range of -1 to +1, where a value close to +1 indicates that the expert agrees that the item is very important to the validity of the content. After CVR value of each item is identified, an item below the value of 0.50 represents questionable validity and is replaced with a new item (Mohamed et al., 2017). The survey question tool gives a CVR value for each item between 0.75 and +1 (**Table 1**) and shows all items that are important and validated and represent statistical thinking processes (Lawshe, 1975).

The reliability of STT was obtained through an agreement between two experts (inter rate-procedure) and evaluators, considering that the study instrument is related to the inconsistency of scoring or individual ranking focus. According to Metwally (2012), the number of inter-raters equal to two people is sufficient for the purposes of this study because the study can use at least two evaluators for the case of obtaining agreement between two experts (inter-rate procedure) and evaluators.

The degree of inter-rater agreement can be assessed by calculating the percentage of agreement between two raters who awarded the same score to students' performance (Reynolds et al., 2008). The inter-rater method used in this study involved two raters, who were excellent mathematics teachers with a background in statistics. These two raters were required to categorize the items in STT, determine the scores for them, and rank the statistical thinking level based on the provided scoring rubric.

Analysis of Data

Mooney's (2002) structure for statistical thinking was used to look at the test results of statistical thought. For each question on the test, first the predicted answers from students were figured out using statistical thought-level descriptions. Two raters appointed amongst the experts in the field of statistics then used this coding method to code student answers to each question on their own in which 88.0% of the time, the writers' codes were the same. There were different codes discussed, and both writers decided on a scale. This rubric was used to code and grade the students' answers. "Demonstrates little awareness of display features, not able to recognize or uses irrelevant features or reasons to recognize the same data represented by different data displays" was part of the unique student answer to the process of defining data. The student whose answer was like this got one point for that question. When a level 4 analytical student answered about the data process, they said that the student "demonstrates complete awareness of display features, including which features are relevant or irrelevant, and uses quantitative relationships between displays to recognize when different displays represent the same data." In this question, the student answer that looked like this adjective got four points. By adding up all of the students' points, we were able to determine their final score on the statistical thought test (Altaylar & Kazak, 2021).

Subsequently, the research data were statistically analyzed using the statistical package for social sciences version 23 software through independent t-test analysis. Before the data is analyzed, the assumption of the t test must first be observed, i.e., a randomly selected study sample, the data converge normally, and the data for each group is homogeneous. The researchers conducted normality to comply with the assumptions of the t-test statistical technique used. In this study, skewness and kurtosis values were used to detect the normality characteristics of each measurement of the variables studied. The ratio values of skewness (-1.28) and kurtosis (-1.21) showed that the data were normally distributed. Data are normally distributed if the kurtosis and skewness values are ±1.96 (Kim, 2013).

RESULTS

Each student's level of thinking was assessed. For each subtask skill, the four levels of statistical thinking (idiosyncratic, transitional, quantitative, and analytical) were assigned numerical values of one, two, three, and four, respectively. The mean scores for each individual and each statistical skill were calculated by collecting the degrees of the subtasks for each statistical skill and dividing them by the number of subtasks (questions in that skill) for each individual separately and for all the study sample members combined. The students were classified according to their statistical thinking levels based on Nemrawi et al. (2022).

The results reveal that students' performance in analyzing and interpreting data is less than satisfactory (mean [M]=1.94, standard deviation [SD]=0.76477). Indeed, students are at a transitional stage at which they can shift between subjective and quantitative evaluations, partially identify units of data value, group or arrange data, but not in a summative manner, and make partially correct comparisons based on one dimension. The findings also indicate that the students showed a higher level of thinking for the described data (M=2.83, SD=0.56806) that formed part of the quantitative level. At this stage, students can organize data, perform systematic data-handling processes, and grasp presentation components, including tables, charts, and graphs, in a summative manner. They are also capable of identifying data units using central tendency and range measures; however, their rationale needs to be stronger, and they depend on quantitative arguments. The value of the standard deviation indicates that the data from the process of describing, organizing and reducing the data are clustered tightly around the mean.

Table 2 presents the descriptive statistics for each item of STT, which shows that the students were at level

Table 2. Distribution of responses

Item	Statistical thinking process	Level of statistical thinking (%)			
		Idiosyncratic	Transitional	Quantitative	Analytical
D1	Describing data	0.0	48.6	51.4	0.0
D2	Describing data	5.7	34.3	48.6	11.4
D3	Describing data	20.0	42.9	28.6	8.6
D4	Describing data	14.3	34.3	40.0	11.4
D5	Describing data	11.4	37.1	40.0	11.4
D6	Describing data	20.0	17.1	48.6	14.3
O7	Organizing and reducing data	14.3	40.0	34.3	11.4
O8	Organizing and reducing data	5.7	40.0	37.1	17.1
O9	Organizing and reducing data	20.0	20.0	34.3	25.7
O10	Organizing and reducing data	11.4	22.9	54.3	11.4
O11	Organizing and reducing data	2.9	37.1	54.3	5.7
R12	Representing data	25.7	34.3	34.3	5.7
R13	Representing data	0.0	91.4	8.6	0.0
R14	Representing data	17.1	40.0	42.0	0.0
R15	Representing data	22.9	51.4	25.7	0.0
R16	Representing data	31.4	25.7	40.0	2.9
R17	Representing data	0.0	60.0	40.0	0.0
A18	Analyzing and interpreting data	60.0	34.3	5.7	0.0
A19	Analyzing and interpreting data	60.0	34.3	5.7	0.0
A20	Analyzing and interpreting data	37.1	57.1	5.7	0.0
A21	Analyzing and interpreting data	25.7	42.9	31.4	0.0
A22	Analyzing and interpreting data	40.0	40.0	20.0	0.0
A23	Analyzing and interpreting data	37.1	40.0	22.9	0.0
A24	Analyzing and interpreting data	28.6	45.7	25.7	0.0
A25	Analyzing and interpreting data	31.4	40.0	28.6	0.0
A26	Analyzing and interpreting data	34.3	37.1	28.6	0.0
A27	Analyzing and interpreting data	25.7	45.7	25.7	2.9
A28	Analyzing and interpreting data	42.9	25.7	31.4	0.0

1 (idiosyncratic), level 2 (transitional), and level 3 (quantitative) for all the items in the analyzing and interpreting data process; however, 32.0% of the students were at level 3 (quantitative; item A28), while only 2.9% of the students were at level 4 (analytical; item A27).

Regarding organizing and reducing the data process, the analysis found that fewer than 30.0% of the students were at the analytical level for all the items. In comparison, the highest percentage, over 50.0% of students, was found at levels 2 (transitional) and 3 (quantitative) for all the items in the organizing and reducing data process. For students at level 2 (transitional) and level 3 (quantitative), more than 65.0% fulfilled all the items in the representing data process. However, only 2.9% of students were at the analytical level for item R16. For the describing data process, the analysis ascertained that more than 70.0% of the students were at level 2 (transitional) and level 3 (quantitative) for all the items. The analysis also found that, for the analyzing and interpreting data process, almost no students reached the analytical level.

Table 3 shows the computed levels of overall statistical thinking on the accumulated statistical skills.

Table 3. Level of statistical thinking (%)

Level	Frequency (n)	Percentage (n)
Idiosyncratic	6	17.1
Transitional	16	45.7
Quantitative	13	37.1
Analytical	0	0.0
Total	35	100

Table 4. Independent sample t-test

Gender	n	M	SD	t value	p-value
Male	12	2.23	0.50518	1.012	0.319
Female	23	2.05	0.52034		

To answer research question 2, the results presented in **Table 3** are used to identify the differences in students' statistical thinking levels based on gender.

An independent t-test was performed (**Table 4**). Levene's test to identify the homogeneity of variance showed $p>0.05$. Thus, the two groups have the same variance value. The result of the t-test analysis was taken from equal variances assumed. The t-test analysis for statistical thinking skills score showed that there was no significant difference in the mean score between the male group ($M=2.23$, $SD=0.50518$) and the female group ($M=2.05$, $SD=0.52034$) with $t(33)=1.012$, $p=.319$. The results therefore revealed that the two groups had equal statistical thinking skills and performance ability. The male and female students were of the same age group; there was a degree of similarity in cognitive processes among younger individuals; and their brains tended to light up similarly. The magnitude of the differences in means (mean difference=.1857, 95% CI=-.1877 to -.5591) was very small (eta squared=.003) (Cohen, 1960). The value of the standard deviation showed that the level of thinking of male is slightly clustered around the mean as compared to female.

DISCUSSION

Based on the information acquired from **Table 3**, it can be seen that most students are at the transitional level. Students' ability to comprehend data presentations might be seen as the foundation for them to begin generating predictions and detecting patterns. At this time, the students are able to demonstrate display feature awareness, recognize the same data in various data displays, assess the efficacy of data displays in expressing data, and identify units of data values. Students can also explain the distribution of the data using fictitious but somewhat valid measurements and create a display that is either full and inaccurately represents the data or half complete and accurate. At the same time, they can assess the efficacy of data visualization by considering pertinent display attributes. Most of the students can make one accurate comparison between data displays or data sets, or a series of somewhat accurate comparisons (Langrall & Mooney, 2002)

According to the study's results in **Table 3**, none of the students attained the level of statistical and analytical thinking. This result implies that students are unable to employ both analytical and numerical answers when dealing with data. The students were incapable of making reasonable inferences based on data and context using various perspectives and reasonably using quantitative relative thinking. By contrast, they were able to achieve the analytical level at which they were able to complete all processes without errors; they could thoroughly understand the data, accurately calculate and link them, explain the purpose of utilizing multiple data presentations and the transition between them, draw relevant conclusions, and generalize from the data (Altaylar & Kazak, 2021; Langrall & Mooney, 2002). The classification of each person for each subtask was unloaded, followed by a compilation of all the classifications based on the type of statistical talent and all the statistical skills to categorize the students according to the degree of their general statistical thinking using pre-prepared forms.

The students' inability to reach the fourth (analytical) level may be attributed to their struggle to construct appropriate statistical arguments as well as their lack of experience in communicating in the statistical language, either verbally or in writing, which is caused by a lack of statistical background, language barriers, or disagreement with statistical concepts, consistent with Woodard et al. (2020). According to Maryati and Priatna (2018), students' poor statistical inference skills may prevent them from reaching the analytical level. Statistical inference uses data analysis to make conclusions about a population or process beyond the available data. Students still have unfavorable views about statistics as a topic as they believe that statistics involves only the ability to compute and employ formulas. The survey also revealed that most pupils had frequent statistical misunderstandings and only possessed a basic understanding of statistics (Nemrawi et al., 2022).

The process of organizing and reducing data has the largest proportion of students at the analytical level. This implies that students may summarize data by forming new categories or clusters, describe data using a valid and accurate measure of center, and describe data spread using a valid and correct measure of spread. The survey also revealed that the majority of students are unable to use quantitative reasoning in the realm of data analysis and interpretation. They are also unable to draw sensible conclusions from the evidence and context.

Additionally, the lack of statistical thinking may be attributed to various conceptual errors in statistics, the most significant of which is the failure to recognize correlation from causality and to differentiate between correlation and mean scores (Nemrawi et al., 2022). The researchers also argued that the low level of statistical thinking could be attributed to the learners' limited understanding of the value of statistical thinking in daily life in many countries as well as the fact that the teaching of statistics is primarily focused on equations and laws and rarely covers its practical applications (Nemrawi et al., 2022). Hence, educating students and teachers about statistics is essential given the importance of statistical thinking in many facets of daily life.

This investigation's outcomes agree with a few other conclusions in the relevant literature (Altaylar & Kazak, 2021). Students frequently showed a level of idiosyncratic statistical thinking; therefore, students need help in establishing statistical thinking. It is thus crucial to concentrate on this throughout their education. Additionally, since defining the data process requires the use of proportional reasoning, as was the case in the investigations by both Jones et al. (2001) and Mooney (2002), the students in this study frequently needed help with this task. Thus, teachers should include proportional thinking tasks in statistical problem settings (Altaylar & Kazak, 2021).

According to the researchers, Piaget's theory of cognitive development, which defined thinking as a chain of synaptic connections and invisible cognitive activities involving the conscious use of the brain to understand the world and choose how to respond to it, can be used to explain the convergence of statistical thinking levels between male and female students (Ghanem, 2017). The lack of disparities in their degrees of statistical thinking is due to the same level of physical development, mental maturity, and prior experience in statistical thinking (Nemrawi et al., 2022). Findings are consistent with those reported by previous researchers (Khamis, 2015; Ramey, 2015) in that no gender differences were found in statistical thinking skills.

Potential discrepancies in achievements between male and female students have long been a subject of interest in statistical studies. However, the influence of gender on the degree of statistical reasoning remains debatable. According to Martin et al. (2017), gender influences the degree of statistical reasoning among third-grade pupils. Indeed, Martin et al. (2017) showed evidence of a gender gap in statistical reasoning favoring men, which is related to differences in anxiety and confidence levels between men and women. Also, when studying gender and statistical thinking, multiple factors should be considered, such as individuals' level of interest in statistical thinking, cognitive processes, and socialization as well as their beliefs, attitudes, motives, and expectations, all of which have a profound influence on individuals' levels of statistical thinking (Nemrawi et al., 2022).

CONCLUSIONS

According to the findings and analysis in this paper, most of the participants in the study sample were at the transitional stage, although some achieved the analytical level. The findings also showed no statistically significant difference in students' statistical thinking degrees based on gender. This study provides recommendations based on the findings to encourage students to develop statistical thinking abilities by comprehending, organizing, condensing, and presenting data in various ways; educators should offer them the opportunity to analyze, interpret, and justify data and encourage verbal or written communication between students through enrichment activities that help them to clarify their thoughts.

Considering the current inadequate level of statistical thinking among students, it is imperative to pursue various initiatives aimed at improving all facets of statistics education, including knowledge, practical application, and abilities. These initiatives include the development of new data exploration tools, improvements in technology usage, and a better understanding of the repercussions and benefits of enhancing statistical thinking and reasoning (Garfield et al., 2015). These tools and improvements in technology use are all part of these efforts. According to Smith et al. (2019) and Watson et al. (2020), when students utilize the STEM framework, they are taught how the data they collect may be evaluated more extensively and methodically to deepen their understanding of the setting under investigation. This allows pupils to have a better understanding of the subject matter under study. This model also provides participants with familiarity with the numerous problem-solving approaches employed in the STEM subjects' various subfields. Students must use statistics because the investigative process in each discipline provides both the context and the variety necessary for such analysis. As a consequence of applying this process, students are better able to recognize obstacles or problems that exist in the real world, collect data, illustrate, analyze, evaluate, and make conclusions (Smith et al., 2019; Watson et al., 2020).

Recommendations

Some suggestions for future researchers and practitioners include conducting qualitative studies by interviewing students and teachers. The study needs to be complementary to determine whether the problem for students and teachers is to foster statistical thinking as well as to examine the components of statistical thought that need more attention. In addition, advanced studies can be carried out by investigating larger samples and various educational arguments to identify significant gaps in the study. The findings found that the majority of students still had a low level of statistical thinking and were unable to analyze and interpret data properly. Some suggestions need to be adjusted so that more research is being carried out to see how students implement each statistical thinking process at various levels of education, especially at the secondary school level.

Limitations

The main limitation of this survey is the small number of respondents. The examination is also limited by a number of time constraints in each course. However, this research has a positive impact on the integration and application of statistical teaching and learning, allowing teachers to design teaching strategies to nurture and enhance students' statistical thinking.

Author contributions: All authors have sufficiently contributed to the study and agreed with the results and conclusions.

Funding: This research was funded by the University Kebangsaan Malaysia (UKM) under grant GUP-2021-028.

Ethical statement: The authors stated the study received an approval from the Ministry of Education (KPM.600-3/2/3-eras (13069), and consent of state education department (JPNS.SPS.MTE-500-12/4 Jld.6 (32) and school administration, instructors, and students to gather the data. Data for this study was gathered using questionnaires. Answers to the survey were anonymous. Further reducing ethical issues, respondents' confidentiality and privacy were safeguarded by keeping their identities anonymous and safely preserving survey data. The study complied with the highest ethical standards, according to the authors. After receiving a thorough explanation of the goals and possible consequences of the study, participants voluntarily gave their informed permission.

Declaration of interest: No conflict of interest is declared by authors.

Data sharing statement: Data supporting the findings and conclusions are available upon request from the corresponding author.

REFERENCES

Altaylar, B., & Kazak, S. (2021). The effect of realistic mathematics education on sixth grade students' statistical thinking. *Acta Didactica Napocensia* [*Napocensia Didactic Act*], *14*(1), 76-90. https://doi.org/10.24193/adn.14.1.6

Aristovnik, A., Kerzic, D., Ravselj, D., Tomazevic, N., & Umek, L. (2020). Impacts of the COVID-19 pandemic on life of higher education students: A global perspective. *Sustainability, 12*(20), 8438. https://doi.org/10.3390/su12208438

Bromage, A., Pierce, S., Reader, T., & Compton, L. (2022). Teaching statistics to non-specialists: Challenges and strategies for success. *Journal of Further and Higher Education, 46*(1), 46-61. https://doi.org/10.1080/0309877X.2021.1879744

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20*(1), 37-46. https://doi.org/10.1177/001316446000200104

Conway, B., Martin, W. G., Strutchens, M., Kraska, M., & Huang, H. (2019). The statistical reasoning learning environment: A comparison of students' statistical reasoning ability. *Journal of Statistics Education, 27*(3), 171-187. https://doi.org/10.1080/10691898.2019.1647008

English, L. D., & Watson, J. (2018). Modelling with authentic data in sixth grade. *International Journal on Mathematics Education, 50*, 103-115. https://doi.org/10.1007/s11858-017-0896-y

Ganesan, N., & Leong, K. E. (2020). Impact of fathom on statistical reasoning among upper secondary students. *Journal of Research in Science, Mathematics and Technology Education,* 3(2), 35-50. https://doi.org/10.31756/jrsmte.321

Garfield, J., Le, L., Zieffler, A., & Ben-Zvi, D. (2015). Developing students' reasoning about samples and sampling variability as a path to expert statistical thinking. *Educational Studies in Mathematics, 88*(3), 327-342. https://doi.org/10.1007/s10649-014-9541-7

Ghanem, B. M. (2017). *Introduction to teaching thinking*. House of Culture for Publishing and Distribution.

Hill, R. (1998). What sample size is "enough" in internet survey research. *Interpersonal Computing and Technology: An Electronic Journal for the 21st Century,* 6(3-4), 1-12.

Jones, G. A., Langrall, C. W., Thornton, C. A., Mooney, E. S., Wares, A., Jones, M. R., Perry, B., Putt, I. J., & Nisbet, S. (2001). Using students' statistical thinking to inform instruction. *Journal of Mathematical Behavior, 20*(1), 109-144. https://doi.org/10.1016/S0732-3123(01)00064-5

Kazak, S., Pratt, D., & Gokce, R. (2018). Sixth grade students' emerging practices of data modelling. *ZDM Mathematics Education, 50*(7), 1151-1163. https://doi.org/10.1007/s11858-018-0988-3

Khamis, R. T. Y. (2015). *The effect of a training program to develop statistical thinking skills among university students* [PhD thesis, University of Baghdad].

Kim, H. Y. (2013). Statistical notes for clinical researchers: assessing normal distribution (2) using skewness and kurtosis. *Restorative Dentistry & Endodontics, 38*(1), 52-54. https://doi.org/10.5395/rde.2013.38.1.52

Langrall, C. W., & Mooney, E. S. (2002). *The development of a framework characterizing middle school students' statistical thinking*. https://api.semanticscholar.org/CorpusID:1588925

Lavidas, K., Barkatsas, T., Manesis, D., & Gialamas, V. (2020). A structural equation model investigating the impact of tertiary students' attitudes toward statistics, perceived competence at mathematics, and engagement on statistics performance. *Statistics Education Research Journal, 19*(2), 27-41. https://doi.org/10.52041/serj.v19i2.108

Lawshe, C. H. (1975). A quantitative approach to content validity. *Personnel Psychology, 28*(4), 563-575. https://doi.org/10.1111/j.1744-6570.1975.tb01393.x

Le, L. (2017). *Assessing the development of students' statistical thinking: An exploratory study* [PhD thesis, University of Minnesota].

Levpuscek, M. P., & Cukon, M. (2022). That old devil called 'statistics': Statistics anxiety in university students and related factors. *Center for Educational Policy Studies Journal,* 12(1), 147-168. https://doi.org/10.26529/cepsj.826

Martin, N., Hughes, J., & Fugelsang, J. (2017). The roles of experience, gender, and individual differences in statistical reasoning. *Statistics Education Research Journal, 16*(2), 454-475. https://doi.org/10.52041/serj.v16i2.201

Maryati, I., & Priatna, N. (2018). Analysis of statistical misconception in terms of statistical reasoning. *Journal of Physics: Conference Series*, *1013*, 012206. https://doi.org/10.1088/1742-6596/1013/1/012206

Matore, M. E. E. M., Khairani, A. Z., Idris, H., & Rahman, N. A. (2017). Kesahan kandungan pakar instrumen IKBAR bagi pengukuran AQ menggunakan nisbah kesahan kandungan [Expert content validity of the IKBAR instrument for AQ measurement using the content validity ratio]. In *Proceedings of the*

International Conference on Global Education V (pp. 979-997).

Metwally, E. (2012). Survey research methods. *Journal of Organizational Change Management, 25*(1), 186-188. https://doi.org/10.1108/0953481121119965

Mohamed, Z., Lebar, O., & Shamsuddin, S. (2017). Pembinaan dan penilaian instrumen ujian aptitud kemasukan ke institut pengajian tinggi Malaysia [Construction and evaluation of aptitude test instruments for admission to Malaysian institutes of higher education]. *Journal of Science and Mathematics Letters, 5*, 16-27. https://doi.org/10.37134/jsml.vol5.2.2017

Mooney, E. S. (2002). A Framework for characterizing middle school students' statistical thinking. *Mathematical Thinking and Learning, 4*(1), 23-63. https://doi.org/10.1207/s15327833mtl0401_2

Nemrawi, A. M., Baioumy, N. A., Fouad, S. A., & Salah. (2022). Statistical thinking levels of students in Al-Ghad International College for applied medical sciences in Saudi Arabia and its relationship to gender. *International Journal of Academic Research in Progressive Education and Development, 1*(2), 398-418. https://doi.org/10.6007/IJARPED/v11-i2/13194

Ramey, J. M. (2015). *Differences in statistical reasoning abilities through behavioral-cognitive combinations of videos and formative assessments in undergraduate statistics courses* [PhD thesis, East Tennessee State University].

Reynolds, C. R., Livingston, R., & Willson, V. (2008). *Measurement and assessment in education*. Pearson.

Rohana, R., & Ningsih, Y. L. (2019). Students' statistical reasoning in statistics method course. *Jurnal Pendidikan Matematika* [*Journal of Mathematics Education*], *14*(1), 81-90. https://doi.org/10.22342/jpm.14.1.6732.81-90

Saidi, S. S., & Siew, N. M. (2022). Assessing secondary school students' statistical reasoning, attitudes towards statistics, and statistics anxiety. *Statistics Education Research Journal, 21*(1), 6. https://doi.org/10.52041/serj.v21i1.67

Schleicher, A. (2022). *Building on COVID-19's innovation momentum for digital, inclusive education*. OECD Publishing. https://doi.org/10.1787/24202496-en

Setambah, M. A. B., Tajuddin, N. M., Yaakob, M. F. M., & Saad, M. I. M. (2019). Adventure learning in basics statistics: Impact on students critical thinking. *International Journal of Instruction, 12*(3), 151-166. https://doi.org/10.29333/iji.2019.12310a

Smith, C., Fitzallen, N., Watson, J., & Wright, S. (2019). The practice of statistics for STEM: Primary students and pre-service primary teachers exploring variation in seed dispersal. *Teaching Science, 65*(1), 38-47.

Syed Zamri, S. N. A., Hutkemri, & Kwan Eu, L. (2020). Postgraduate students' attitude toward statistics pre and post scenario-based learning method in a statistics course. *Malaysian Online Journal of Educational Sciences, 8*(4), 1-8.

United Nations. (2021). *The sustainable development goals report 2021*. https://unstats.un.org/sdgs/report/2021/The-Sustainable-Development-Goals-Report-2021.pdf

Watson, J. M. (2013). Resampling with *Tinker*Plots. *Teaching Statistics, 35*(1), 32-36. https://doi.org/10.1111/j.1467-9639.2012.00511.x

Watson, J., Fitzallen, N., English, L., & Wright, S. (2020). Introducing statistical variation in year 3 in a STEM context: Manufacturing licorice. *International Journal of Mathematical Education in Science and Technology, 51*(3), 354-387. https://doi.org/10.1080/0020739X.2018.1562117

Woodard, V., Lee, H., & Woodard, R. (2020). Writing assignments to assess statistical thinking. *Journal of Statistics Education, 28*(1), 32-44. https://doi.org/10.1080/10691898.2019.1696257

World Bank. (2022). *Helping countries adapt to a changing world*. https://www.worldbank.org/en/about/annual-report

World Economic Forum. (2020). *The future of jobs report 2020*. https://www.weforum.org/publications/the-future-of-jobs-report-2020/

Yusoff, M. S. B. (2019). ABC of content validation and content validity index calculation. *Education in Medicine Journal, 11*(2), 49-54. https://doi.org/10.21315/eimj2019.11.2.6