MODESTUM EURASIA Journal of Mathematics, Science and Technology Education, 2024, 20(12), em2552
ISSN:1305-8223 (online)
OPEN ACCESS Review Article https://doi.org/10.29333/ejmste/15708

What are statisticians contemplating?–A thirty-year literature review and trend analysis on statistical thinking

Wen-Mao Chung [1,2], Hsin-Chueh Chen [3,4], Mei-Chun Lin [1,5], Pei-Chi Chen [1,6], Chun-Yen Chang [1,7*]

[1] National Taiwan Normal University, Taipei, TAIWAN
[2] Renai Elementary School, Tainan City, TAIWAN
[3] National Taiwan University of Science and Technology, Taipei, TAIWAN
[4] University of Twente, Enschede, THE NETHERLANDS
[5] Shulin Elementary School, New Taipei City, TAIWAN
[6] New Taipei City Shulin High School, New Taipei City, TAIWAN
[7] Universitas Negeri Malang, Malang, INDONESIA

Received 03 October 2023 ▪ Accepted 10 March 2024

Abstract

The term ‘statistical thinking’ might seem straightforward at face value, yet defining it precisely remains a challenge. What exactly constitutes statistical thinking? And how does it differentiate from other forms of thinking? To date, the academic community lacks a universally accepted definition. Thus, this study employed a retrospective research method, examining and analyzing the discourse on statistical thinking in international journals over the past 30 years, integrating findings from 44 papers. Beyond analyzing research trends, this study also compiled seven definitions or structures of statistical thinking. The results indicate that various fields have favored definitions or structures, but regardless of the chosen definition or structure, there is a consistent association with the concept of variation. Concluding, the researchers reviewed the curriculum guidelines for 12-year basic education and the implications of statistical education within Taiwan’s curriculum, offering tangible recommendations for future investigations.

Keywords: literature review, mathematics education, statistical education, statistical thinking

INTRODUCTION

Throughout our lives, a multitude of information or data relies on statistics for analysis or forecasting, examples being the likelihood of rain in weather reports, price oscillations, referendum polls, and satisfaction surveys. Given the inherent variability in such collected data, the application of statistics becomes indispensable. Statistical procedures encompass several stages: initial contemplation, problem identification, planning of statistical endeavors, data gathering, processing, scrutiny, and, ultimately, conclusion drawing. Such conclusions might either align with the initial hypotheses or raise new queries, prompting further statistical undertakings.

Traditional statistics education has chiefly emphasized the significance of mathematical and logical thinking (Su Guoliang, 1999). While mathematical thinking focuses on simplifying problems and establishing mathematical models, logical thinking centers on deducing cause-and-effect relationships and reasoning. However, Shaughnessy and Ciancetta (2002) noted that current statistical teaching seems to overly highlight the mathematical aspect, overlooking the distinctiveness of statistical thinking. Statistical thinking not only pertains to data analysis but also encompasses thinking about the interconnections between data and associated phenomena (Su Guoliang, 1999). In other words, it melds mathematics and logic, but with a stronger emphasis on the dynamic nature and variability of data. Chen Xingmei (2006) further asserted that the heart of statistics should incorporate statistical thinking, which fundamentally serves as the cornerstone of the entire statistical process.

✉ amoun0501@hotmail.com ✉ lilychen@mail.ntust.edu.tw ✉ fresh438@gmail.com ✉ peggy89330537@gmail.com
✉ changcy@ntnu.edu.tw **(*Correspondence)**

Contribution to the literature

- It conducts a thorough literature review on statistical thinking over the past 30 years, highlighting evolving academic perspectives.
- It identifies seven diverse definitions and structures of statistical thinking, enriching the academic dialogue.
- It examines and provides recommendations for statistical education in Taiwan, emphasizing the importance of developing statistical thinking in education.

Wilks (1951) stated in his American Statistical Association address that one day, statistical thinking would be as essential to an effective citizen as foundational skills like reading and writing. Nonetheless, there exists a diversity of definitions for statistical thinking amongst scholars, leaving its precise definition somewhat fragmented and unresolved.

This study, therefore, reviews literature spanning the past thirty years related to statistical thinking, integrating the various definitions or structures proposed by different academics. In the process of data encoding, this research draws upon Laudan's (1984) reticulated model of science as a theoretical foundation, emphasizing the interdependencies, mutual constraints, and coordination among methods, subjects, and theories. This tripartite framework guides our integration and analysis of the literature. Despite the ambiguity surrounding the definition of statistical thinking, statistics remain prominently utilized within the mathematical community. Thus, this research further delves into the prevailing trends in statistical research amongst mathematicians. Based on the findings, the researchers offer recommendations for the mathematics curriculum.

Research Questions

Drawing upon a comprehensive literature review and analysis, the researchers collated and scrutinized journal articles pertaining to "statistical thinking", aiming to address the subsequent queries:

1. How is "statistical thinking" defined in literature spanning the past thirty years? What frameworks for statistical thinking have been conceived or employed?
2. What trends can be discerned within the literature regarding statistical thinking over the last three decades?
3. How might one enhance students' capacity for statistical thinking?

METHODS

Literature Collection Methods

For this study, we utilized the Academic Search Premier comprehensive subject full-text database (hosted by EBSCOhost) for our literature search. To ensure a thorough collection, databases such as Education Research Complete, MathSciNet via EBSCOhost, Professional Development Collection, Education Resource Information Center, and Teacher Reference Center were explored, employing "statistical thinking" as the keyword for searches, whilst restricting the type of document to "journal articles". Following this procedure, we gathered a total of 44 papers, from which one non-English article and two off-topic articles were excluded. The reasons for excluding these two off-topic articles were, as follows: one was an editorial column, which leaned more towards the author's personal opinions rather than academic research; the other discussed the concept of six standard deviations, a topic that, while related to statistics, did not align with the specific scope of our study.

The journals of these 44 papers were categorized using the Web of Science (http://mjl.clarivate.com/). Out of these, 19 papers were found in SCI/SCIE/SSCI, representing 43.2%; 15 were indexed in the Emerging Sources Citation Index (ESCI), representing 34.1%; and 10 were located in other journals, comprising 22.7% of the total. Nonetheless, exclusion from the citation index database does not diminish a journal's significance. For instance, the Mathematics Teacher is a publication by the National Council of Teachers of Mathematics. This council notably released the "Principles and Standards for School Mathematics", which possesses a global influence. Therefore, even if journals are not indexed within SCI/SCIE/SSCI/ESCI, the articles they feature remain valuable for analysis. The journals chosen by the researchers for these papers are presented in **Table 1**.

Literature Analysis Framework & Coding Methods

Whilst the authors identified several articles related to "statistical thinking" during their database search, a systematic review on the topic remained absent. Hence, they opted to utilize Laudan's (1984) "reticulated model of scientific rationality" as the analytical framework. This model underscores the importance of not solely focusing on individual research methods, topics, or theories, but rather on their interconnections. Such an approach ensures a holistic exploration of "statistical thinking". In terms of research methodology, we were inspired by the classifications introduced by Wang Wenke and Wang Zhihong (2019), which include survey, observational, experimental, case study, and literary

Table 1. Classification of journals for selected literature

Journal	Database	n	P (%)	Reference
The American Statistician	SCI/SCIE	4	9.1	Chick and Watson (2002), Coleman (2013), Scariano and Calzada (2007), & Tong (2019)
Journal of Mathematical Behavior	ESCI	2	4.5	Capraro et al. (2005) & Hoerl et al. (2014)
Educational Studies in Mathematics	SSCI	1	2.3	Groth (2005)
The Mathematics Teacher	-	3	6.8	Groth and Powell (2004), Korakianiti and Rekkas (2011), & Scheaffer (2002)
Teaching Statistics	ESCI	5	11.4	Garfield et al. (2015), Jones et al. (2000), White (2015), Wild and Pfannkuch (1999), & Zhang and Stephens (2016)
PNA	ESCI	1	2.3	Aizikovitsh-Udi et al. (2014)
Elementary Education Online	-	1	2.3	Vandenbroeck et al. (2006)
OD Practitioner	-	1	2.3	Sedlmeier (2000)
Total Quality Management	SSCI	3	6.8	Bajaria (1997), Garfield and DelMas (2010), & Steel et al. (2019)
International Journal of Mathematical Education in Science & Technology	ESCI	1	2.3	Peled et al. (2013)
Instructional Science	SSCI	1	2.3	Shaughnessy and Pfannkuch (2002)
Interdisciplinary Description of Complex Systems.	ESCI	1	2.3	Bjerke (2002)
Journal of Biopharmaceutical Statistics	SCI	1	2.3	Velury (1997)
WIREs: Computational Statistics	ESCI	1	2.3	Koparan and Guven (2013)
Teaching of Psychology	SSCI	1	2.3	Snee (1990)
IIE Solutions	-	1	2.3	Viles (2008)
Pharmaceutical Research	SCI/SCIE	1	2.3	Lancaster (2011)
Journal of College Science Teaching	-	1	2.3	Lee (2015)
Mathematical Thinking and Learning	SSCI	2	4.5	Derry et al. (1995) & Groth (2006)
Annals of the Institute of Statistical Mathematics	SCIE	1	2.3	Akaike (2010)
International Statistical Review	SCI/SCIE	2	4.5	Schwinn (2017) & Woltman (2017)
Korean Journal of Anesthesiology	ESCI	1	2.3	Melton (2004)
Statistical Journal of the IAOS	-	1	2.3	Mooney (2002)
The Mathematics Enthusiast	ESCI	1	2.3	Boland (2003)
Journal of the American Statistical Association	SCI/SCIE	1	2.3	Schuyten and Thas (2007)
Managerial Auditing Journal	SSCI	1	2.3	Mooney (2002)
Journal of STEM Education	ESCI	1	2.3	English and Watson (2015)
Eurasia Journal of Mathematics, Science & Technology Education	-	1	2.3	Žmuk (2015)
School Science and Mathematics	-	1	2.3	Burgess (2009)
Journal of Statistics Education	ESCI	1	2.3	Kugler et al. (2003)

Note. n: Number of articles & P: Percentage

analysis approaches. The authors conducted a qualitative analysis coding manually after reading each paper individually. This manual coding process ensured the depth and quality of the research, with the coded categorization results detailed in tables.

When it comes to research topics, given that no existing literature review framework specific to statistical thinking is available, we carried out an in-depth analysis of the literature, leading to the categorization, as follows:

1. **Improvement of product yield and efficiency in quality management:** This reflects statistical thinking's role in enhancing production and quality, aiming to boost efficiency.
2. **Suggestions on research methodologies:** Highlights the importance of statistical thinking in developing robust research designs, improving research reliability.
3. **Development of a conceptual module framework:** Emphasizes the need for structured tools to aid in understanding and applying statistical concepts.
4. **Recommendations for statistical teaching:** Focuses on teaching strategies to foster students' critical thinking and problem-solving through statistics.
5. **Curriculum development and design:** Centers on curriculum design to enhance statistical thinking, providing a structured approach for educators.

These categories showcase the significance of statistical thinking in various aspects, from education to practical application, guiding comprehensive understanding and targeted recommendations for its integration and development across fields. For topics 3, 4, and 5, a more detailed categorization was applied, differentiating based on the subjects of research–notably distinguishing among teachers, university students, secondary school students, and primary school pupils. Concerning research theory, given the diverse perspectives from different disciplines and authors on the definition and framework of statistical thinking, we identified and categorized seven distinct definitions and frameworks for statistical thinking to aid coding. The process of article coding was undertaken by a primary school teacher and a secondary school teacher, both of whom are doctoral students in science education. To enhance the reliability among coders, several measures were implemented. First, a clear set of coding rules and guidelines was established to minimize the impact of subjective judgments.

Additionally, a pilot test was conducted prior to the main coding phase to verify the effectiveness of the coding framework and make necessary adjustments. Throughout the coding process, regular reviews were performed to check for consistency in coding, and adjustments were made to the coding framework as needed. In instances, where discrepancies arose during coding, discussions were held until a consensus was reached. The coding framework can be found in **Table 2**.

RESULTS & DISCUSSION

Definition & Framework of Statistical Thinking

The uniqueness of statistical thinking lies in its integrative approach, encompassing not only the comprehension and manipulation of data but also the critical evaluation of the assumptions, conditions, and limitations underlying the data.

Table 2. Coding scheme for literature analysis

Item	Analysis code	Reference
Research methods	Experimental methods	Peled et al. (2013) & Shaughnessy and Pfannkuch (2002)
	Observational methods	Akaike (2010), Garfield and DelMas (2010), Garfield et al. (2015), Lancaster (2011), Melton (2004), Scariano and Calzada (2007), Schuyten and Thas (2007), Tong (2019), Viles (2008), & Zhang and Stephens (2016)
	Survey	Bjerke (2002), & Žmuk (2015)
	Literary analysis	Bajaria (1997), Chick and Watson (2002), Groth (2005), Koparan and Guven (2013), Makrymichalos et al. (2005), Mooney (2002), Schwinn (2017), Sedlmeier (2000), & Steel et al. (2019)
	Case study	Aizikovitsh-Udi et al. (2014), Boland (2003), Burgess (2009), Capraro et al. (2005), Coleman (2013), Derry et al. (1995), English and Watson (2015), Groth (2006), Groth and Powell (2004), Hoerl et al. (2014), Jones et al. (2000), Korakianiti and Rekkas (2011), Kugler et al. (2003), Lee (2015), Scheaffer (2002), Snee (1990), Vandenbroeck et al. (2006), Velury (1997), White (2015), Wild and Pfannkuch (1999), & Woltman (2017)
Topics	Improvement of product yield & efficiency in quality management	Bajaria (1997), Bjerke (2002), Garfield and DelMas (2010), Lancaster (2011), Mooney (2002), Sedlmeier (2000), Steel et al. (2019), Velury (1997), & Viles (2008)
	Suggestions on research methodologies	Akaike (2010), Chick and Watson (2002), Koparan and Guven (2013), Melton (2004), & Schuyten and Thas (2007)
	Development of a conceptual module framework	Derry et al. (1995), Groth (2006), Tong (2019), & Woltman (2017)
	Recommendations for statistical teaching	Boland (2003), Burgess (2009), Capraro et al. (2005), Garfield et al. (2015), Groth (2005), Hoerl et al. (2014), Jones et al. (2000), Kugler et al. (2003), Lee (2015), Makrymichalos et al. (2005), Scariano and Calzada (2007), Schwinn (2017), Vandenbroeck et al. (2006), & Žmuk (2015)
	Curriculum development & design	Aizikovitsh-Udi et al. (2014), Coleman (2013), English and Watson (2015), Groth and Powell (2004), Korakianiti and Rekkas (2011), Peled et al. (2013), Scheaffer (2002), Shaughnessy and Pfannkuch (2002), Snee (1990), White (2015), Wild and Pfannkuch (1999), & Zhang and Stephens (2016)
Research objects	Teachers	Aizikovitsh-Udi et al. (2014), Boland (2003), Snee (1990), & Žmuk (2015)
	University students	Coleman (2013), Groth (2005), Lee (2015), Peled et al. (2013), Shaughnessy and Pfannkuch (2002), White (2015), Wild and Pfannkuch (1999), Woltman (2017), & Žmuk (2015)
	Secondary school students	Derry et al. (1995), Groth and Powell (2004), Hoerl et al. (2014), Jones et al. (2000), Korakianiti and Rekkas (2011), Scariano and Calzada (2007), Scheaffer (2002), Vandenbroeck et al. (2006), & Zhang and Stephens (2016)
	Primary school students	Burgess (2009), Capraro et al. (2005), English and Watson (2015), Groth (2006), & Zhang and Stephens (2016)

Table 2 (Continued). Coding scheme for literature analysis

Item	Analysis code	Reference
Research theory	Framework: Statistical thinking in quality improvement (Snee, 1990)	Garfield and DelMas (2010), Mooney (2002), & Tong (2019)*
	Framework: An empirical framework for statistical thinking (Wild & Pfannkuch, 1999)	Boland (2003), Garfield et al. (2015), Groth (2005), Korakianiti and Rekkas (2011), Schwinn (2017), Woltman (2017)*
	Framework: Representational framework for children's statistical thinking, mathematical thinking, & learning (Jones et al., 2000)	Groth (2006)*
	Framework: Representational framework for secondary school students' statistical thinking, mathematical thinking, & learning (Mooney, 2002)	Derry et al. (1995)* & Vandenbroeck et al. (2006)
	Definition: Mallows (1998)	Chick and Watson (2002)
	Definition: American Society for Quality (ASQ, 1996)	Bjerke (2002), Koparan and Guven (2013), Kugler et al. (2003), Lancaster (2011), Sedlmeier (2000), & Steel et al. (2019)
	Definition: Hoerl and Snee (2002)	White (2015)

Note: Entries marked with * denote that framework is original

As Moore (1990) highlighted, statistical thinking involves a comprehensive process from identifying issues to collecting and analyzing data, and finally making reasoned inferences, necessitating both technical skills and a high level of critical thinking and problem-solving abilities. The essence of statistical education/thinking is aimed at developing a set of skills and knowledge that enhance an individual's ability to understand and analyze data using statistical methods. This includes traditional statistical knowledge and skills such as data collection, analysis, and interpretation, as well as the application of this knowledge in making decisions in daily life.

Key attributes of statistical thinking involve sensitivity to data, understanding of variability and uncertainty, and the capacity to assess the reliability of analytical outcomes. The background and characteristics of statistical education encompass data awareness, acknowledgment of inherent variability and uncertainty in data analysis, mastery of various statistical tools and techniques for data interpretation, problem-solving using statistical approaches, and critical evaluation of analytical assumptions, methods, and conclusions. Furthermore, the interdisciplinary nature of statistics, with its application across a wide range of fields such as social sciences, natural sciences, medicine, and engineering, underscores the need for statistical education to not only focus on statistical knowledge and skills but also to equip educators and learners with the ability to adapt and think across disciplines. This prepares them to tackle specific challenges and problems in various fields.

In this section, we will delve deeply into the definition and framework of 'statistical thinking'. Particular emphasis will be placed on introducing the four distinct frameworks proposed by Jones et al. (2000), Mooney (2002), Snee (1990), and Wild and Pfannkuch (1999). Detailed descriptions of each are provided, as follows:

Snee (1990). Statistical thinking in quality improvement

Snee (1990) perceives statistical thinking as a cognitive process that emerges amidst variability, a form of fluctuation pervasive in every aspect of life. Every endeavor constitutes an interconnected systemic process. Thus, it's imperative to identify, characterize, quantify, and manage these variations to elucidate a definitive path for improvement.

Wild and Pfannkuch (1999). An empirical study on statistical thinking

In their 1999 study, Wild and Pfannkuch (1999), through interviews with statisticians, introduced a structured framework for statistical thinking. This framework offers a comprehensive and organized approach to understanding and applying statistical thinking. The framework comprises the following four primary components:

1. **Investigative cycle–**This section outlines the entire procedure of statistical research, from the identification of problems to the derivation of conclusions.
2. **Interrogative cycle–**This segment characterizes the cognitive processes statisticians engage in when faced with challenges and data. It encompasses the generation of query loops, searching, interpretation, critique, and judgement.
3. **Types of thinking–**This zeroes in on the five main modes of statistical thinking. It spans recognizing the importance of data, reasoning with models,

and integrating both background and statistical knowledge.

4. **Dispositions–**This pertains to the individual traits that come to the fore during the thinking process, potentially influencing the style and result of statistical analyses.

Jones et al. (2000). Representational frameworks for children's statistical, mathematical thinking, & learning

In their 2000 study, Jones et al. (2000), drawing upon the structure of the observed learning outcome (SOLO) taxonomy proposed by Biggs and Collis in 1982 and 1991, presented a cognitive development framework for statistical thinking tailored for students from the first to the fifth grades. They delved deep into the essence of statistical thinking, segmenting it into four primary tiers, which collectively form the central framework of statistical thinking. These four tiers are:

1. **Idiosyncratic:** At this level, students' thinking is unique, largely shaped by individual experiences.
2. **Transitional:** Here, students begin to establish connections and comprehend the foundational characteristics of data.
3. **Quantitative:** The thinking becomes more profound at this stage, with students capable of conducting quantitative analysis of data.
4. **Analytic:** Students are equipped to undertake in-depth analyses, interpret data, and derive conclusions from it.

Mooney (2002). Representational architectures for statistical thinking, mathematical thinking, & learning in middle school students

In designing a framework for statistical thinking, Mooney also referenced SOLO taxonomy presented by Biggs and Collis in 1982 and 1991. He proposed a cognitive development structure for statistical thinking for students in years six to eight. This framework parallels that introduced by Jones et al. (2000) in its form, though the content at each level varies based on different themes.

Beyond these four definitions and frameworks for statistical thinking, we have further distilled general definitions of statistical thinking from other sources and will elaborate in detail in the subsequent sections.

Mallows (1998)

The definition of statistical thinking underscores the close connection between data and the real world, specifically focusing on variability and uncertainty, aiming to delve into and comprehend the genuine implications concealed within the data.

American Society for Quality

Statistical thinking is a philosophy of learning and action based on three core principles:

(1) all work occurs within a system of interconnected processes;

(2) there is inherent variation in every process; and

(3) understanding and controlling this variation are essential for success (American Society for Quality [ASQ], 1996).

Hoerl and Snee (2002)

Statistical thinking amalgamates the scientific method with the notion of variation, aiming to delve deeper into understanding the processes and the data these processes generate.

Surveying the previous seven points, it becomes clear that, apart from the third and fourth points, which primarily rely on SOLO taxonomy, the other five perspectives, whether in terms of a framework or a definition, predominantly highlight the central theme of 'variation'. This suggests that variation holds paramount importance in statistical thinking. Put succinctly, the true essence of statistical thinking revolves around understanding and analyzing "variation".

Identifying Preferred Definition or Framework of Statistical Thinking in Each Research Field

In this study, we explicitly delineated seven distinct definitions and frameworks of statistical thinking. When comparing these definitions and frameworks across different research areas, we found that the definition proposed by ASQ (1996) was the most favored among researchers and was cited in six papers. Notably, this is the only definition of statistical thinking that is recognized across disciplines. Given that the main focus of the database in this study is the educational field, we speculate that if the research scope were broader, ASQ (1996) definition would likely be cited even more frequently. The wide acceptance of this definition among researchers may be attributed to its clear and comprehensive description, as well as the authoritative position of the organization ASQ (1996) as an internationally renowned body. Following closely is the statistical thinking framework proposed by Wild and Pfannkuch (1999), which is also cited in six papers. However, this framework is primarily applied to statistical research activities, and as such, it is mostly used in papers related to education. All research results are summarized and presented in **Table 3**.

In educational academic papers, the most frequently cited is the statistical thinking framework by Wild and Pfannkuch (1999), closely followed by the definition given by ASQ (1996). Within the realm of management sciences, the frameworks utilized predominantly derive from established frameworks within the management

Table 3. Allocation of statistical thinking's definition/structure across various fields

Statistical thinking definition/architecture	Education (n=35)	Management science (n=7)	Pharmaceutical industry (n=2)	Subtotal
1. Snee (1990)	1	2	0	3
2. Wild and Pfannkuch (1999)	6	0	0	6
3. Jones et al. (2000)	1	0	0	1
4. Mooney (2002)	2	0	0	2
5. Mallows (1998)	1	0	0	1
6. ASQ (1996)	2	3	1	6
7. Hoerl and Snee (2002)	1	0	0	1
Not explicitly mentioned	21	2	1	24

Table 4. Allocation of articles without a clear definition of statistical thinking across various citation index databases

Citation index database	Number of articles lacking clear definition	Total number of papers reviewed	Proportion (%)
SCI/SCIE/SSCI	8	19	42.0
ESCI	9	15	60.0
None	7	10	70.0
Subtotal	24	44	54.5

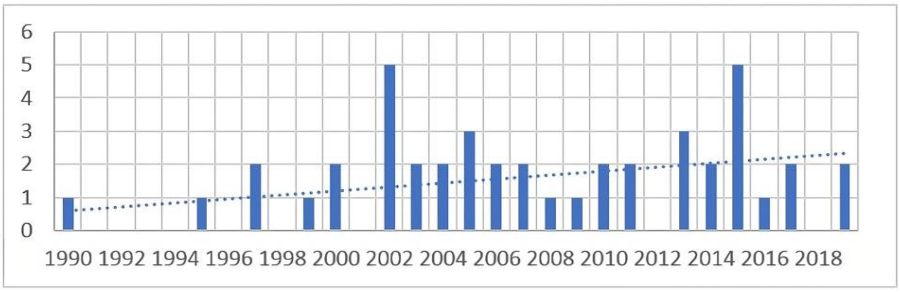

Figure 1. Number of published papers from 1990 to 2019 (Source: Authors' own elaboration)

science community. Specifically, ASQ (1996) is cited in three articles, while Snee (1990) is referenced in two.

From **Table 3**, it is evident that out of the 44 articles discussing statistical thinking, 24 do not clearly define statistical thinking. This suggests that a significant portion of these publications lacks a clear definition of statistical thinking.

When cross-referencing these 24 articles with their respective citation index databases, it is found that among articles classified under SCI/SCIE/SSCI, the proportion without a clear definition is the lowest at 42.0%. For ESCI, this figure stands at 60.0%. Notably, for articles not indexed within any database, the proportion lacking a clear definition rises to a striking 70.0%. Detailed data can be seen in **Table 4**.

Analysis of Trends in Statistical Thinking Research

Time distribution

The earliest article in this study was published in 1990, with the most recent one appearing in 2019, spanning approximately three decades. The data reveals that the second article emerged six years after the initial 1990 publication. Furthermore, out of the 44 articles, 39 were published post-2000, constituting nearly 90.0% of the total. Since the turn of the millennium, there has been an average annual publication rate of about two articles. The annual publication count is illustrated in **Figure 1**.

As depicted in **Figure 1**, since the start of 2000, there's been a significant uptick in the number of papers published, with pronounced peaks in 2002 and 2015. The decade-long interval between these peaks may reflect the academic community's in-depth exploration and continuous discussions on the precise definition of 'statistical thinking'. While debates and varied perspectives persisted, the ubiquity of statistics ensured undying enthusiasm amongst researchers. Overall, this trend underscores the escalating recognition and appreciation of statistical thinking.

Country distribution

Out of the 44 papers, 22 hail from the United States, making up 50.0% of the total. The United Kingdom stands as the second most prolific contributor. Belgium, New Zealand, and Australia each have contributed two papers. 11 countries, including Germany, Japan, Norway, Ireland, Greece, Croatia, South Korea, Israel, Spain, and China, have penned just one article each. Put succinctly, the top-five countries together account for three-quarters, or 75.0%, of the total count. The overarching distribution is depicted in **Figure 2**.

Figure 2 highlights that the predominant countries are all developed nations with notably high human

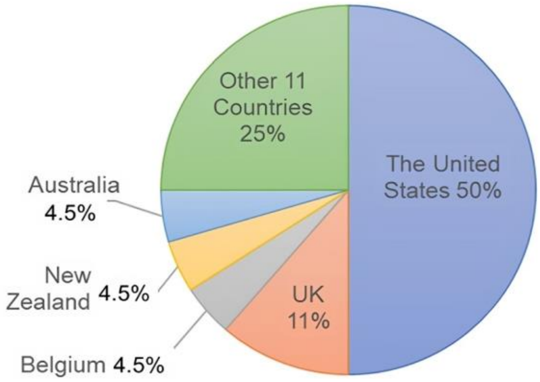

Figure 2. Pie chart of proportion of papers (Source: Authors' own elaboration)

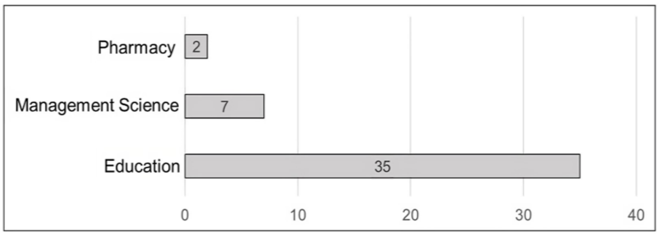

Figure 3. Distribution of research fields (Source: Authors' own elaboration)

Table 5. List of papers in citation index databases

Country/database	SCI/SCIE/SSCI	ESCI	Other	Total
United States	9	6	7	22
United Kingdom	2	2	1	5
Belgium	2	0	0	2
New Zealand	1	1	0	2
Australia	0	2	0	2
Germany	1	0	0	1
Japan	1	0	0	1
Norway	1	0	0	1
Ireland	1	0	0	1
Greece	1	0	0	1
Croatia	0	1	0	1
South Korea	0	1	0	1
Israel	0	1	0	1
Spain	0	1	0	1
China	0	0	1	1
Turkey	0	0	1	1
Total	19	15	10	44

development indices. When juxtaposing the countries with the three types of citation index databases, namely SCI/SCIE/SSCI, it emerges that of the 19 papers, nine are from the United States, whilst the United Kingdom and Belgium each contributed two papers, collectively making up nearly 70.0% of the submissions. New Zealand, Germany, Japan, Norway, Ireland, and Greece each have one paper, with the detailed breakdown provided in **Table 5**.

Western nations, especially those in Europe and the United States, are considerably ahead in publishing papers on statistical thinking when compared to Eastern nations. Within Eastern countries, only Japan, South Korea, and China have published papers, with a combined total of three, representing less than 10.0%. One might speculate that this could be due to language barriers, but upon closer examination, this does not appear to be the primary factor. The top five countries globally in terms of populations, where English is an official language include India, the United States, Pakistan, Nigeria, and the Philippines; with the United States being an exception, the majority are situated in the Asian region. As such, the researchers propose that this significant disparity might arise from a combination of the level of national economic development and the influence of the official language.

Research Fields & Topics

Considering the research fields, those related to education boast the highest number of papers, amounting to 35, which represents approximately 79.5% of the total papers. This is followed by management science, with seven papers making up about 16.0% of the overall count, and two in the pharmaceutical sector, constituting roughly 4.5% of the total, as illustrated in **Figure 3**.

The pharmaceutical industry seeks to enhance the quality and efficacy of its products. With regard to quality management improvement, to maintain a leading position, it's essential not only to advance and mature in technology but also to prioritize the enhancement of production quality (Korakianiti & Rekkas, 2011). Consequently, the pharmaceutical sector utilizes statistical methods to boost its production efficiency and quality, gives attention to matters of statistical thinking, and undertakes academic exchanges by organizing pertinent seminars (Vandenbroeck et al., 2006).

Management science is a key discipline that engages with statistical thinking. Among the papers collated and analyzed in this study, the earliest one was authored by Snee (1990), a research authority in management science. Snee (1990) posits that for statistics to be relevant in the context of total quality management (TQM), they need to diverge from traditional models. The epitome of TQM is not achieved through statistical tools, but rather through statistical thinking (Snee, 1990).Whilst statistical tools focus on data analysis, statistical thinking emphasizes the broader context and meaning behind the data. In TQM, this translates to looking beyond mere figures to deeply understand the overall quality management process. Since the advent of Snee's (1990) work in 1990, it seems to have heralded the study of statistical thinking. This domain applies the statistical thinking paradigm of management science, and numerous related research have been robustly pursued in diverse sectors.

However, the exploration and advocacy of statistical thinking inevitably circle back to the domain of

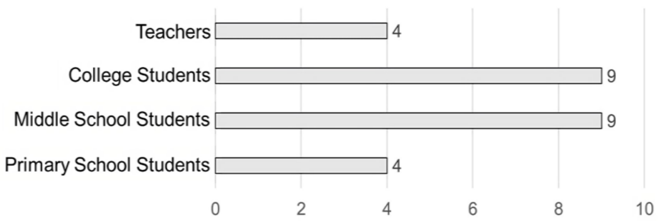

Figure 4. Distribution of research subjects within educational domains (Source: Authors' own elaboration)

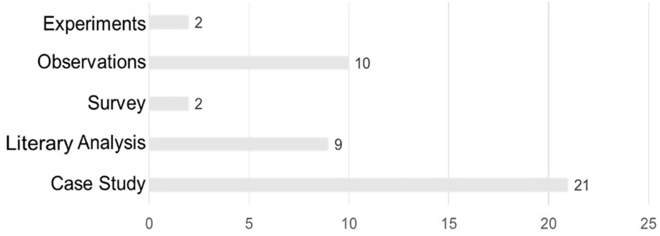

Figure 5. Distribution of research methodology (Source: Authors' own elaboration)

education. The majority of papers analyzed in this study pertain to education. After excluding nine literary analysis papers, the research subjects of the remaining 26 papers were analyzed. It emerges that middle school students and university students are predominant, with nine papers each, together comprising 70.0% of the educational papers. In contrast, primary school pupils and educators are represented in four papers each, equating to 30.0% of the educational papers, as depicted in **Figure 4**.

The authors of these papers are often affiliated with universities or conducting studies within university settings. It's hardly surprising then that university students, being the most accessible research subjects, feature so predominantly. Middle school students, too, are a popular choice amongst researchers. They often present more potential than primary school pupils, especially in terms of communicative abilities and foundational statistical knowledge. Conversely, there are fewer papers, where teachers and primary school students are the focal subjects. This observation could serve as a pointer for directions in future research.

Research Methodology

Out of the 44 articles, case studies lead the way with 21 papers, making up about 48.0% of the total. Many of these papers targeted specific educational groups for their investigations. Coming next, observational research is present in 10 articles, equating to roughly 23.0% of the corpus. These typically interweave the author's individual experiences with theoretical frameworks for comprehensive analysis. Literary analyses are also abundant, featuring in nine articles, amounting to approximately 20.0% of the total. The detailed discussion on these analyses is given before. Strikingly, experimental and survey methods, which are staples in generic research, find themselves less prevalent in studies around statistical thinking, each appearing in only two articles. Collectively, they comprise about 9.0% of the dataset. The abstract nature of 'thinking' perhaps renders direct documentation challenging, prompting researchers to veer towards qualitative methodologies, favoring in-depth interviews or observational approaches. The research methodology utilized in the studies under scrutiny are presented in **Figure 5**.

In the domain of statistical thinking research, there's a noted preference for utilizing existing large datasets over employing traditional experimental or survey research methodologies. This trend can be attributed to several factors. Firstly, leveraging existing data allows for the examination of a broad range of hypotheses without incurring the significant costs associated with new data collection, thereby enhancing research efficiency and optimizing resource utilization (Berman, 2018). Furthermore, advancements in statistical and machine learning technologies have simplified the extraction of meaningful insights from existing datasets, reducing the necessity for collecting new data (James et al., 2013). Moreover, statistical thinking research emphasizes the identification of genuine patterns within data, necessitating stringent bias control and the enhancement of study external validity. Traditional experimental and survey methods may fall short in controlling certain biases, affecting the reliability of research outcomes (Dunning, 2012). Although experimental research methods can provide evidence of causality, they face practical challenges in design and implementation, such as ensuring experimental control and sample representativeness. Similarly, survey methods are susceptible to response and non-response biases, limiting their applicability in statistical thinking studies (Groves et al., 2009; Shadish et al., 2002). Hence, while experimental and survey methodologies retain their value in specific contexts, their application in the realm of statistical thinking is comparatively limited. Conversely, observational study methods, content analysis, and case study approaches are extensively utilized within statistical thinking research due to their ability to delve into the complexity of phenomena and provide a rich source of data. These methodologies facilitate a deeper understanding of the multifaceted interactions within social sciences and offer a wealth of qualitative and quantitative data, enabling researchers to thoroughly analyze issues and gain a better understanding of individuals' attitudes, values, and motivations. The selection of these methods reflects an acknowledgment of the unique needs within the statistical education field and fosters a deeper exploration of the connections between result interpretation and the distinctive characteristics of the statistical education domain.

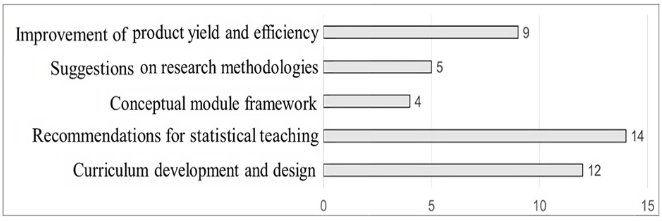

Figure 6. Topic distributions within research (Source: Authors' own elaboration)

Research Topics

The breadth of research topics has been categorized into five distinct topics: Improvement of product yield and efficiency in quality management; proposing suggestions on research methodologies; development of a conceptual module framework furnishing recommendations for statistical teaching; and sculpting curriculum development and design. **Figure 6** shows topic distributions within research.

Of the 44 papers scrutinized in this study, 35 pertain to the realm of education, constituting a substantial portion of the aggregate. Consequently, it is evident that "proffering recommendations for statistical teaching" is the predominant topic, covered in 14 papers; "curriculum development and design" is the subject of 12 papers. Furthermore, there are 9 papers dedicated to "Improvement of product yield and efficiency in quality management" and five pieces centered on "proposing suggestions on research methodologies."

Several studies emphasized the need to enhance students' statistical thinking, offering a variety of **pedagogical recommendations**. Boland (2003) proposed using real-life scenarios from students' daily experiences to make statistical concepts more relatable and understandable, thus bridging the gap between abstract statistics and everyday decision-making. Kugler et al. (2003) highlighted the necessity of fundamental statistical concepts in biostatistics education, reflecting the interdisciplinary nature of statistics and its critical role in scientific research. This approach underscored the importance of a solid statistical foundation for effective data analysis across various scientific fields. Melton (2004) discussed the crucial role of data collection in statistics education, emphasizing that understanding and engaging in continuous data collection can develop students' ability to derive insights from data, highlighting statistics' practical aspect. Sedlmeier (2000) recommended an interactive and visual learning approach, which helped students grasp abstract statistical concepts more effectively by applying them in tangible settings, thereby enhancing statistical thinking through active engagement. Schuyten and Thas (2007) advocated for the use of interactive multimedia techniques to make statistical learning more dynamic and engaging, thereby improving students' learning outcomes in statistics through the application of modern technology. Garfield and DelMas (2010) introduced the ARITIST (assessment resource tools for improving statistical thinking) website for resource sharing, promoting the development of statistical thinking across different educational stages by providing extensive educational resources, thus fostering a community of learning. Lastly, Chick and Watson (2002) examined the impact of cooperative learning in statistics, highlighting how teamwork and social interaction could deepen students' understanding of statistical concepts and improve their problem-solving abilities, reflecting the collaborative nature of statistical analysis. These diverse recommendations collectively aimed to enhance statistical thinking among students, aligning with the unique interpretative and applied nature of statistical education.

In **curriculum development and design**, past studies have shown that students' statistical inference abilities are often derived from real-life experiences. Shaughnessy and Pfannkuch (2002) incorporated natural phenomena like Old Faithful's eruption patterns into their curriculum to link statistical education with practical applications, helping students see the relevance of statistics in natural and scientific contexts. In contrast, Stephen and Maria (2007) engaged students in finding personally relevant data online, arguing that discovering compelling data encourages genuine statistical thought, transitioning students from passive observers to active participants. Recognizing the gap in statistical inquiry skills among adults, Derry et al. (1995) suggested using real-world problems as educational resources to enhance students' analytical skills regarding societal issues, aiming to foster practical skills and social responsibility. Viles (2008) followed a similar approach by creating exercises that simulate real-life scenarios, improving students' problem-solving and inferential reasoning. Woltman (2017) connected life safety education with statistical teaching using public road casualty data, aiming to make students more aware and reduce vehicle accident risks among teenagers. This approach not only boosts statistical understanding but also addresses societal safety concerns. Finally, Aizikovitsh-Udi et al. (2014) developed a model to foster critical and statistical thinking, suggesting these skills are interdependent. White (2015) used Excel for practical exercises, promoting critical thinking within statistical tasks. Steel et al. (2019) critiqued standard statistics courses for overemphasizing computational errors, advocating instead for fostering sound reasoning to understand and correct misconceptions, highlighting the necessity of integrating real-life context and critical analysis into statistical education.

In the domains of managerial science and pharmaceuticals, statistical thinking plays an indispensable role in **elevating product yield and driving excellence in quality management**. Bajaria (1997) argues that statistics must encompass knowledge from various domains and apply statistical methods and techniques in a multidisciplinary context to enhance the quality of products and services. Integrating expertise from different scientific fields to optimize product and service quality is essential. Furthermore, emphasizing the importance of statistics in the decision-making process for individuals, large corporations, or government institutions is vital to minimize risks and prevent errors (Schwinn, 2017), highlighting the significance of statistical thinking in risk management and decision-making capabilities. Schwinn further incorporates this statistical perspective into organizational development strategies, underlining its key role in quality management, a necessity even for small businesses to support their growth and competitiveness. However, many small enterprises, due to resource constraints, neglect investing in statistical training for their employees and essential tools. In fact, a survey by Žmuk (2015) in Croatia reveals that proper utilization of statistical techniques could significantly improve the performance of over 90.0% of small businesses. Regrettably, up to 93.0% of such companies have not integrated statistics into their operational strategies, a neglect that could undeniably limit their growth potential. In facing the challenges of the new era, Snee (1990) emphasized that American companies have begun to recalibrate their strategies, particularly focusing on improving product yield and the efficiency of quality management. This overall enhancement in standards, termed 'total quality,' is crucial for American businesses to maintain a competitive edge in fierce international competition. Snee further explained that without a comprehensive understanding and application of statistical thinking in quality management, navigating the rapid changes in the global market would be challenging. At that time, although discussions on statistical thinking were prevalent, clear definitions and understanding were scarce. To bridge this gap, Snee presented a definition of statistical thinking, illustrating its crucial role at every stage of 'total quality.'

Addressing **academic research guidelines**, Tong (2019) firmly believes that meticulous and systematic planning is indispensable during the research design phase. Furthermore, one should employ inferential statistics with utmost caution. Historical scientific breakthroughs, such as Kepler's laws, the periodic table, the germ theory of disease, DNA's structure, and energy quantification, all emerged without resorting to inferential statistics. The author argues that guiding statistical thinking, along with prudent utilization of inferential statistics, can mitigate concealed errors often inherent in data analytics and inferential methodologies. Regarding statistics, it inherently involves mastering data manipulation, hence the fervent discussions around big data. As cited in the research by Hoerl et al. (2014), WIRED's editor-in-chief published an article in 2008 titled "The end of theory; the abundance of data makes the scientific method obsolete", sparking widespread attention in the scientific community. When pondering the nuances of big data, it's undeniably possible to derive conclusions by amalgamating extensive datasets and swift computations powered by cutting-edge technology. Nonetheless, the author contends that beyond mere data crunching and discourse, proficient modelling and statistical thinking are pivotal for big data's success. He underscores the significance of offering suggestions on research methodologies, emphasizing the inextricable link between statistical thinking and its practical application in research.

In contrast to application and advocacy, only four papers delve deeply into the **development of the conceptual module framework** designed for the application domains and subjects of "statistical thinking". One paper is excluded, as it is not relevant to the development of a framework for statistical thinking. The detailed exploration of these papers is given before, where their contributions to the conceptual framework for statistical thinking are thoroughly discussed.

"Variation" as Cornerstone of Statistical Thinking

Watson and Kelly (2002) posit that data variation lies at the heart of statistics; without variation in data, there's no need for statistics. In the context of statistical thinking, Moore (1990) had previously highlighted that variation is its core component. He noted, "variation pervades the entire process ... One must account for variation when devising data output or collection methods ... Concerning the interpretation of variation ...". Similarly, Wild and Pfannkuch (1999) concurred, placing emphasis on variation as a crucial element in statistical thinking.

Upon revisiting the seven frameworks or definitions of statistical thinking compiled in this study, one can observe congruent insights. After excluding Jones et al. (2000) and Mooney (2002)–who designed a hierarchical framework for assessing the statistical thinking of primary and secondary students without explicitly mentioning the relationship between variation and statistical thinking–the keyword "variation" is evident in the remaining five definitions or frameworks, as detailed in **Table 6**.

Reading and Reid (2004) also believed that the success of a statistics course largely hinges on whether the course design integrates considerations of variation factors into classroom statistical activities. However, for younger students, they can be directed to observe the phenomenon of variation, employing informal reasoning alongside the situation to instruct them on how to embrace and manage variation.

Table 6. "Statistical thinking" definition or framework & "variation"

Definition or framework of statistical thinking	"Variation" mentioned
Framework: Statistical thinking in quality improvement (Snee, 1990)	...**Variations** are all around us… We must identify, characterize, quantify, & control these **variations** to offer opportunities for improvement.
Framework: An empirical architecture for statistical thinking (Wild & Pfannkuch, 1999)	Mentioned in second dimension: Consider **variation**; & proposes four points of discussion, including noticing & acknowledging **variation** ... measuring & modelling **variation** ... explaining & handling **variation** ... employing strategies to observe **variation** ...
Definition: Mallows (1998)	Statistical thinking concentrates on link between data itself & real world, particularly regarding **variation** & uncertainty ...
Definition: American Society for Quality (ASQ, 1996)	Statistical thinking is a philosophy of action and study rooted in three principles: (1) All work transpires within an interconnected system of processes; (2) **Variation** is present in all processes;& (3) Comprehending and minimizing **variation** is pivotal to success.
Definition: Hoerl and Snee (2002)	Statistical thinking merges scientific method & notion of **variation** to aid in grasping processes & data these processes generate.

CONCLUSIONS & RECOMMENDATIONS

Definition & Framework of Representative Statistical Thinking

The study scrutinized 44 articles across 30 research journals encompassing education, management science, and the pharmaceutical sector. It identified that the definition of statistical thinking posited by ASQ (1996) enjoys broad adoption across fields like education, management science, and the pharmaceutical industry, marking it as the prevailing definition of statistical thinking across diverse sectors. Regarding the framework of statistical thinking, this study discerned that the four-dimensional framework proposed by Wild and Pfannkuch (1999), rooted in empirical research, is most favored in educational research. It's predominantly employed to depict various states during statistical undertakings.

In light of the statistical thinking framework proffered by Wild and Pfannkuch (1999), the primary dimension encompasses the investigation cycle. This cycle involves problem discovery (P), plan formulation (P), data collection (D), data analysis (A), and conclusion presentation (C). Notably, statistical activity is intrinsically an investigative cycle. Furthermore, two additional statistical thinking frameworks utilized in this study's educational research were put forth by Jones et al. (2000) and Mooney (2002). These two frameworks bear significant resemblances and are constructed around statistical activities, incorporating aspects such as data description, organization and data simplification, data presentation, as well as data analysis and interpretation. Comparatively, P-P-D-A-C process outlined by Wild and Pfannkuch (1999) represents a holistic statistical activity procedure. The frameworks from Jones et al. (2000) and Mooney (2002), when set against various PPDAC stages, seem to culminate merely at the data collection (D) and data analysis (A) phases. This arguably mirrors the compartmentalized design of primary and secondary school statistical education, which breaks down the statistical process. By merely teaching students about diverse statistical graphs or concepts, without imparting the rationale for data collection (P) or the planning behind it (P), potential biases in subsequent statistical outcomes and conclusion presentations (C) may arise.

Nevertheless, the frameworks proposed by Jones et al. (2000) and Mooney (2002) offer hierarchical distinctions across data processing stages (level 1-level 4). These can be employed to evaluate students' levels of statistical thinking, an aspect that was absent in the architecture of Wild and Pfannkuch (1999). Each framework for statistical thinking boasts its distinct features, leading researchers to select the one most apt for their specific research topics and subjects.

Trends in Research on Statistical Thinking

Despite this study's comprehensive review of literature spanning nearly three decades, we observed a significant concentration of research on statistical thinking primarily post-2000, with nearly 90.0% of studies published in the last two decades. This trend reflects the growing emphasis within the field of statistical education on advancing statistical thinking. Analysis of research domains revealed that, aside from the educational category, which constitutes a significant majority (79.5%), the fields of production and management also demonstrated considerable attention (20.5%). However, within educational research, studies focusing on primary school students and teachers were relatively scarce, and most employed case study methodologies. Empirical research, including experiments and surveys, was relatively rare, accounting for less than 10.0% of studies. This phenomenon may stem from over half of the studies (54.5%) failing to clearly define the concept of statistical thinking, posing challenges in designing specific experiments or surveys and determining what data to collect and how. Moreover, authors may assume familiarity with the concept of statistical thinking among their readers, hence considering it unnecessary to define it explicitly within their articles.

The distinction between statistical thinking and statistical concepts highlights the need for educators to transcend traditional concept teaching in instructional design, focusing more on developing students' cognitive abilities and problem-solving strategies. Although many studies aim to enhance students' statistical thinking, there is often a lack of detailed description in the literature regarding exactly what aspects of thinking were improved and the extent of this improvement. Establishing a grading matrix similar to the one developed by Jones et al. (2000) and Mooney (2002) for assessing students' statistical thinking could significantly facilitate the quantitative evaluation of educational outcomes and enhance teaching quality.

The findings of this study present new challenges and opportunities within the field of statistical education research. For instance, the observation that studies on statistical thinking in primary and secondary education are relatively few and seldom use grading tools for assessing the development of statistical thinking not only highlights potential biases in literature collection and research methodology selection but also underscores the necessity for more specific and targeted research on statistical thinking. Furthermore, as American statistical courses evolve over time, there is a need to update or modify these grading tools to meet contemporary educational demands.

Statistical Thinking in Curriculum Guidelines for 12-Year Basic Education

Variation in statistics is the central component of statistical thinking, primarily because the world is rife with variation, and much is fraught with uncertainty. In the most recent curriculum guidelines for 12-year basic education released by the Ministry of Education in 2019 (corresponding to Taiwan's Minguo year 108), the mathematics syllabus for national primary, secondary, and senior secondary schools is divided into four predominant learning themes: number and quantity (N), space and shape (S), relationship (R), and data and uncertainty (D). It's noteworthy that what was labeled as "statistics and probability" in the 2008 version (Taiwan's Minguo year 97) of the Ministry of Education's curriculum has been renamed to "data and uncertainty" in the 2019 version (Taiwan's Minguo year 108). This shift highlights the increased focus on data processing and the nuances of uncertainty within statistical investigations.

Upon examining the cycle of statistical surveys, the learning content for the first grade at the primary school stage is centered around "simple classification." The second grade delves into "classification and presentation," the third grade explores "one-dimensional tables and two-dimensional tables," the fourth grade emphasizes "registering bar charts and line charts and constructing bar charts," the fifth grade tackles "producing line charts," and the sixth grade is dedicated to "circle charts." A meticulous observation reveals that the course content is primarily restricted to data processing, with the absence of a comprehensive statistical investigation activity. For senior high school, the seventh-grade theme is "statistical charts and statistics," the eighth grade is "statistical data processing," the ninth grade introduces "distribution of statistical data," and the tenth grade emphasizes "data analysis and systematic counting." The curriculum for the eleventh and twelfth grades lacks topics related to statistical activities.

Considering the forthcoming curriculum guidelines for 12-year basic education in mathematics, when paralleling the statistical process with the statistical survey cycle P-P-D-A-C, it becomes evident that the courses designed encompass only D (data collection) and A (data analysis), lacking a holistic curriculum. Solely concentrating on data collection, analysis, or conclusions offers a truncated comprehension of statistical processes (Shaughnessy, 2007). Thus, it remains imperative for mathematics teachers to curate related activities, enabling students to undergo a full-fledged statistical investigation activity.

Future Research Directions

Groth (2006) posited that our society is increasingly reliant on statistical methods and their outcomes. Consequently, mathematics teachers hold a pivotal role in enhancing students' proficiency in statistical thinking. To adequately support students, it's imperative to first comprehend their statistical thinking capabilities. Given the abstract nature of statistical thinking, how should we gauge students' proficiency in this domain? Wild and Pfannkuch (1999) have proffered a thought framework for statistical activities, while Jones et al. (2000) and Mooney (2002) have assessed primary and secondary school students' thinking within D-A phase of P-P-D-A-C cycle. Although P-P and C elements are not formally structured into the primary and secondary curricula in our nation, primary school students might find it challenging to correlate questions, evidence, and conclusions. Nevertheless, with appropriate guidance, executing a full statistical survey activity is within reach for primary school pupils (Fielding-Wells, 2010). As a way forward, there's potential to devise a comprehensive hierarchical framework of statistical thinking, equipping primary and secondary educators to assess students' levels of statistical thinking.

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

Funding: No funding source is reported for this study.

Ethical statement: The authors stated that the study followed ethical standards for social science research. Approval from an ethics committee was not required since it is a review of existing literature.

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

Aizikovitsh-Udi, E., Clarke, D., & Kuntze, S. (2014). Hybrid task design: Connecting learning opportunities related to critical thinking and statistical thinking. *PNA, 8*(3), 115-126. https://doi.org/10.30827/pna.v8i3.6115

Akaike, H. (2010). Making statistical thinking more productive. *Annals of the Institute of Statistical Mathematics, 62*(1), Article 3. https://doi.org/10.1007/s10463-009-0238-0

ASQ. (1996). *Glossary and tables for statistical quality control*. Quality Press.

Bajaria, H. J. (1997). Total quality management and statistical thinking. *Total Quality Management, 8*(243), S1-S4. https://doi.org/10.1080/0954412979631

Ben-Zvi, D., & Friedlander, A. (1997). Statistical thinking in a technological environment. In J. Garfield and G. Burrill (Eds.), *Research on the role of technology in teaching and learning statistics* (pp. 45-55). International Statistical Institute. https://doi.org/10.52041/SRAP.96103

Berman, E. P. (2018). *Creating the market university: How academic science became an economic engine*. Princeton University Press.

Bjerke, F. (2002). Statistical thinking in practice: Handling variability in experimental situations. *Total Quality Management, 13*(7), 1001-1014. https://doi.org/10.1080/0954412022000017076

Boland, P. J. (2003). Promoting statistical thinking among secondary school students in the national context. *The American Statistician, 57*(2), 85-88. https://doi.org/10.1198/0003130031432

Burgess, T. (2009). Teacher knowledge and statistics: What types of knowledge are used in the primary classroom? *The Mathematics Enthusiast, 6*(1), 3-24. https://doi.org/10.54870/1551-3440.1130

Capraro, M. M., Kulm, G., & Capraro, R. M. (2005). Middle grades: Misconceptions in statistical thinking. *School Science and Mathematics, 105*, Article 165. https://doi.org/10.1111/j.1949-8594.2005.tb18156.x

Chen Xingmei. (2006). The connotation and teaching philosophy of statistics course in elementary school. *Science Education Monthly, 287*, 2-12.

Chick, H. L., & Watson, J. M. (2002). Collaborative influences on emergent statistical thinking–A case study. *Journal of Mathematical Behavior, 21*, 371-400. https://doi.org/10.1016/S0732-3123(02)00135-9

Coleman, S. Y. (2013). Statistical thinking in the quality movement ±25 years. *The TQM Journal, 25*(6), 597-605. https://doi.org/10.1108/TQM-06-2013-0075

Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches*. SAGE.

Derry, S., Levin, J. R., & Schauble, L. (1995). Stimulating statistical thinking through situated simulations. *Teaching of Psychology, 22*(1), 51-56. https://doi.org/10.1207/s15328023top2201_16

Dunning, T. (2012). *Natural experiments in the social sciences: A design-based approach*. Cambridge University Press. https://doi.org/10.1017/CBO9781139084444

English, L. D., & Watson, J. M. (2015). Exploring variation in measurement as a foundation for statistical thinking in the elementary school. *Journal of STEM Education, 2*, Article 3. https://doi.org/10.1186/s40594-015-0016-x

Fielding-Wells, J. (2010). Linking problems, conclusions and evidence: Primary students' early experiences of planning statistical investigations. In C. Reading (Ed.), *Proceedings of the 8th International Conference on Teaching Statistics*. International Statistical Institute.

Garfield, J., & DelMas, R. (2010). A web site that provides resources for assessing students' statistical literacy, reasoning and thinking. *Teaching Statistics, 32*, 2-7. https://doi.org/10.1111/j.1467-9639.2009.00373.x

Garfield, J., delMas, R., & Chance, B. (2003). *The web-based ARTIST: Assessment resource tools for improving statistical thinking project* [Paper presentation]. The AERA Annual Meeting.

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*, 327-342. https://doi.org/10.1007/s10649-014-9541-7

Groth, R. E. (2005). An investigation of statistical thinking in two different contexts: Detecting a signal in a noisy process and determining a typical value. *Journal of Mathematical Behavior, 24*, 109-124. https://doi.org/10.1016/j.jmathb.2005.03.002

Groth, R. E. (2006). An exploration of students' statistical thinking. *Teaching Statistics, 28*(1), 17-21. https://doi.org/10.1111/j.1467-9639.2006.00003.x

Groth, R. E., & Powell, N. N. (2004). Using research projects to help develop high school students' statistical thinking. *Mathematics Teacher, 97*, 106-109. https://doi.org/10.5951/MT.97.2.0106

Groves, R. M., Fowler Jr, F. J., Couper, M. P., Lepkowski, J. M., Singer, E., & Tourangeau, R. (2009). *Survey methodology*. Wiley.

Hoerl, R. W., Snee, R. D., & De Veaux, R. D. (2014). Applying statistical thinking to 'big data' problems. *WIREs: Computational Statistics, 6*, 222-232. https://doi.org/10.1002/wics.1306

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning*. Springer. https://doi.org/10.1007/978-1-4614-7138-7

Jones, G. A., Thornton, C. A., Langrall, C. W., Mooney, E. S., Perry, B., & Putt, I. J. (2000). A framework for characterizing children's statistical thinking. *Mathematical Thinking and Learning, 2*(4), 269-307. https://doi.org/10.1207/S15327833MTL0204_3

Koparan, T., & Guven, B. (2013). A study on the differentiation levels of middle school students' statistical thinking. *Elementary Education Online, 12*(1), 158-178.

Korakianiti, E., & Rekkas, D. (2011). Statistical thinking and knowledge management for quality-driven design and manufacturing in pharmaceuticals. *Pharmaceutical Research, 28*(7), 1465-1479. https://doi.org/10.1007/s11095-010-0315-3

Kugler, C., Hagen, J., & Singer, F. (2003). Teaching statistical thinking. *Journal of College Science Teaching, 32*(7), Article 434. https://doi.org/10.2505/4/jcst03_032_07_434

Lancaster, G. A. (2011). How statistical literacy, official statistics and self-directed learning shaped social enquiry in the 19th and early 20th centuries. *Statistical Journal of the IAOS, 27*, 99-111.

Lee, S. (2015). The first step to good clinical research: Statistical thinking. *Korean Journal of Anesthesiology, 68*(2), 99-100. https://doi.org/10.4097/kjae.2015.68.2.99

Makrymichalos, M., Antony, J., Antony, F., & Kumar, M. (2005). Statistical thinking and its role for industrial engineers and managers in the 21st century. *Managerial Auditing Journal, 20*(4), 354-363. https://doi.org/10.1108/02686900510592043

Melton, K. I. (2004). Statistical thinking activities: Some simple exercises with powerful lessons. *Journal of Statistics Education, 12*(2). https://doi.org/10.1080/10691898.2004.11910736

Ministry of Education. (2008). *1997 national primary and secondary school nine-year consistent curriculum–Mathematics learning area*. Ministry of Education.

Ministry of Education. (2018). *Twelve-year national basic education curriculum outline national primary and secondary schools and ordinary senior secondary schools–Mathematics*. Ministry of Education.

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

Moore, D. S. (1990). Uncertainty. In L. Steen (Ed.), *On the shoulders of giants: New approaches to numeracy* (pp. 95-137). National Academy Press.

Neuman, W. L. (2013). *Social research methods: Qualitative and quantitative approaches*. Pearson.

Peled, O. N., Peled, I., & Peled, J. U. (2013). Uncommon knowledge of a common phenomenon: Intuitions and statistical thinking about gender birth ratio. *International Journal of Mathematical Education in Science and Technology, 44*(1), 59-69. https://doi.org/10.1080/0020739X.2012.690898

Reading, C., & Reid, J. (2004). Measuring the development of students' consideration of variation. *Statistics Education Research Journal, 7*(1), 40-59. https://doi.org/10.52041/serj.v7i1.479

Scariano, S. M., & Calzada, M. E. (2007). Statistical thinking with trend charts. *Mathematics Teacher, 100*(8), 534-541. https://doi.org/10.5951/MT.100.8.0534

Scheaffer, R. (2002). Statistical bridges. *Journal of the American Statistical Association, 97*(457), 1-7. https://doi.org/10.1198/016214502753479176

Schuyten, G., & Thas, O. (2007). Statistical thinking in computer-based learning environments. *International Statistical Review, 75(3)*, 365-371. https://doi.org/10.1111/j.1751-5823.2007.00031.x

Schwinn, D. (2017). Statistical thinking for the OD professional. *OD Practitioner, 49(2)*, 47-54.

Sedlmeier, P. (2000). How to improve statistical thinking: Choose the task representation wisely and learn by doing. *Instructional Science, 28*(3), 227-262. https://doi.org/10.1023/A:1003802232617

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Houghton Mifflin.

Shaughnessy, J. M. (2007).Research on statistics learning and reasoning. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning*. National Council of Teachers of Mathematics.

Shaughnessy, J. M., & Ciancetta, M. (2002). Students' understanding of variability in a probability environment. In *Proceedings of the 6th International Conference on Teaching Statistics: Developing a Statistically Literate Society, Cape Town, South Africa*. International Statistical Institute.

Shaughnessy, J. M., & Pfannkuch, M. (2002). How faithful is old faithful. *Mathematics Teacher, 95*(4), 252-259. https://doi.org/10.5951/MT.95.4.0252

Snee, R. D. (1990). Statistical thinking and its contribution to total quality. *The American Statistician, 44*(2), 116-121. https://doi.org/10.2307/2684144

Steel, E. A., Liermann, M., & Guttorp, P. (2019). Beyond calculations: A course in statistical thinking. *The*

American Statistician, 73(sup1), 392-401. https://doi.org/10.1080/00031305.2018.1505657

Su Guoliang. (1999). Enlightenment and development of statistical concepts. *Science Education Monthly, 220*, 9-16.

Tong, C. (2019). Statistical inference enables bad science; Statistical thinking enables good science. *The American Statistician, 73*(sup1), 246-261. https://doi.org/10.1080/00031305.2018.1518264

Vandenbroeck, P., Wouters, L., Molenberghs, G., Van Gestel, J., & Bijnens, L. (2006). Teaching statistical thinking to life scientists a case-based approach. *Journal of Biopharmaceutical Statistics, 16*(1), 61-75. https://doi.org/10.1080/10543400500406520

Velury, J. (1997). Statistical thinking in problem solving. *IIE Solutions, 29*(10), Article 30.

Viles, E. (2008). Paper-clip case: A practical activity to improve statistical thinking for engineering students. *Teaching Statistics, 30*(2), 57-60. https://doi.org/10.1111/j.1467-9639.2008.00322.x

Wang Wenke, & Wang Zhihong (2019). *Educational research law*. Wunan Publishing House.

Watson, J. M., & Kelly, B. A. (2002). Grade 5 students' appreciation of variation. In A. Cockburn, & E. Nardi (Eds.), *Proceedings of the 26th Conference of the International Group for the Psychology of Mathematics Education* (pp. 385-392). University of East Anglia.

White, D. (2015). Active learning and threshold concepts in multiple testing that can further develop student critical statistical thinking. *Teaching Statistics, 37*(2), 48-53. https://doi.org/10.1111/test.12069

Wild, C. J., & Pfannkuch, M. (1999). Statistical thinking in empirical enquiry. *International Statistical Review, 67*(3), 223-265. https://doi.org/10.2307/1403699

Wilks, S. S. (1951). Presidential address to the American Statistical Association. *JASA, 46*(253), 1-18. https://doi.org/10.1080/01621459.1951.10500763

Woltman, M. (2017). Promoting statistical thinking in schools with road injury data. *Teaching Statistics, 39*(1), 26-29. https://doi.org/10.1111/test.12117

Yin, R. K. (2014). *Case study research: Design and methods*. SAGE.

Zhang, Q., & Stephens, M. (2016). Profiling teacher capacity in statistical thinking of national curriculum reform: A comparative study between Australia and China. *EURASIA Journal of Mathematics, Science and Technology Education, 12*(4), 733-746. https://doi.org/10.12973/eurasia.2016.1225a

Žmuk, B. (2015). Business sample survey measurement on statistical thinking and methods adoption: The case of Croatian small enterprises. *Interdisciplinary Description of Complex Systems: INDECS, 13*(1), 154-166. https://doi.org/10.7906/indecs.13.1.14