CHAPTER III
RESEARCH METHODOLOGY

A. Research Method and Design

This research is a cross-national comparative study. The mixed method is used because it enables the researcher to use quantitative and qualitative data to solve problems that are too complicated to address from a single-method standpoint(Creswell, 2009). The quantitative data and their analysis procedures only give a general grasp of the research problem, whereas the qualitative data and their analysis enrich and contextualize those statistical results by scrutinizing participants' perspectives.

The mixed-method strategy used in this study is the sequential explanatory strategy. This two-phase, sequential mixed methods study aims to investigate STEM teachers' teaching beliefs. This research gathered quantitative and qualitative data collection timing in phases or sequentially. The sequential model is shown in the following Figure 4.

This mixed-method design strategy often appeals to researchers with solid quantitative leanings. In this design, the researcher first collects and analyzes numerical data in the quantitative phase. The qualitative phase follows, where text-based data is collected and analyzed to further explain and elaborate on the quantitative results obtained in the first phase. The rationale behind this approach is that the quantitative data and analysis offer a broad understanding of the research problem, while the qualitative data and analysis refine and clarify the statistical results by delving deeper into participants' perspectives.

This research obtained data from a questionnaire for STEM teachers, STEM lesson plans, and teacher interviews. In the first phase, quantitative research questions addressed the teaching beliefs of secondary STEM teachers in Indonesia and Taiwan, which consist of STEM education goals, teaching approaches, self-efficacy, and commitment. Information from this first phase was explored further in a second qualitative phase. In the second phase, qualitative analysis of STEM lesson plans and teacher interviews were used to probe significant teachers' perceptions of STEM teaching by exploring aspects of the teaching beliefs with teachers who agreed to collect their lesson plans and to be interviewees both in Indonesia and Taiwan.

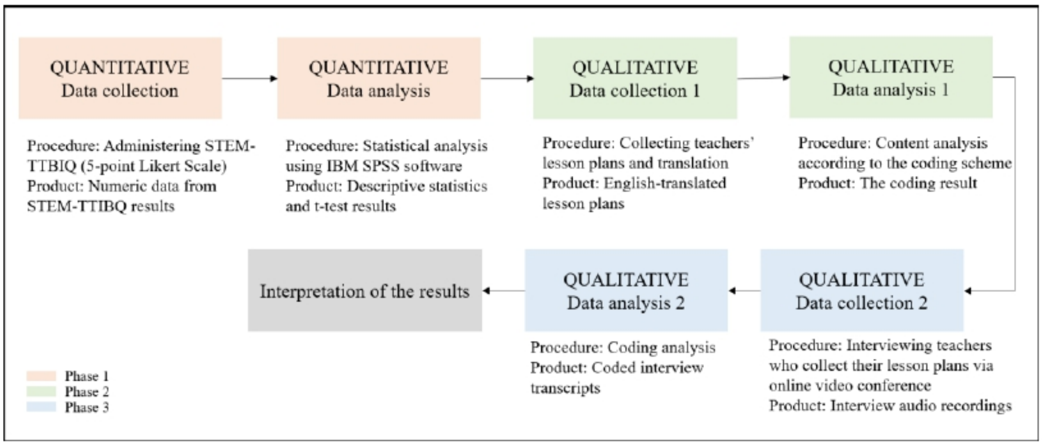

Figure 4. Sequential explanatory mixed method design

This design's advantage includes exploring quantitative results in greater detail. This design proves particularly valuable when unexpected outcomes emerge from a quantitative study. However, the limitations of this design lie in the time-consuming nature of collecting and analyzing both types of data and the resources required to do so effectively (Ivankova et al., 2006). In addition to the research design in Figure 4, a more elaborate research procedure is shown in Figure 5 to show the comparative characteristic of this study.

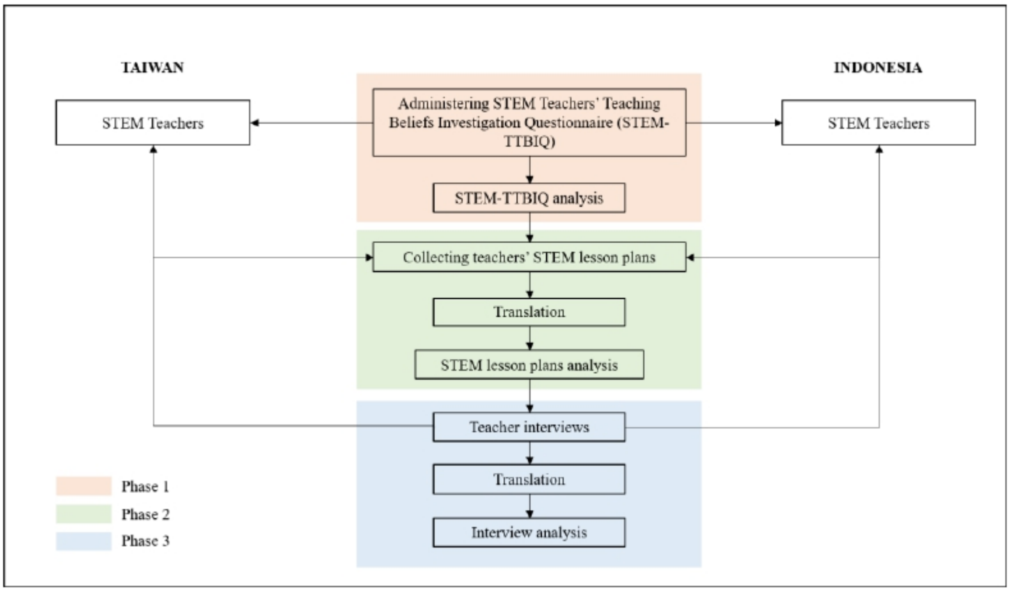

Figure 5. Research Procedure

B. Population and Sample

The population for this study comprises junior high school or middle school teachers who have received STEM teaching professional development training and actively engage in STEM teaching and learning in their classrooms, commonly referred to as pioneer teachers. This criterion applies to participants from both Indonesia and Taiwan. STEM teaching experiences in this context include conducting STEM teaching and learning processes in classrooms or organizing STEM projects in science clubs.

In the case of Taiwan, prior research involved the administration of a quantitative survey. To supplement this, qualitative data collection methods such as lesson plans and interviews are employed alongside the three data collection steps conducted in Indonesia. The selection of secondary schools in Taiwan was carried out using the proportionate stratified sampling method. The sampled schools were then contacted to recruit teachers who teach Science, Mathematics, or Technology-related subjects.

Of the 63 participating Taiwanese secondary schools, 705 teachers voluntarily and anonymously responded to the questionnaire. Participants who left most sub-scales unanswered (more than 50%) or provided meaningless responses were excluded from the dataset. The final sample comprises 653 teachers' responses, with fewer than four missing values per item, reflecting a valid response rate of 92.6%. Among these respondents, 83 teachers reported practicing STEM teaching, while 55 teachers reported organizing STEM activities in their schools.

To ensure comparability, despite potential variations in the number of STEM teachers participating in the study between Indonesia and Taiwan, the sampling technique focuses on equity. Careful consideration is given to the unique qualities and characteristics of teachers from both countries, ensuring that participants possess similar qualifications. As a result, the questionnaire was completed by 69 Taiwanese teachers with STEM teaching experience and 40 Indonesian teachers, maintaining the desired comparability in the study.

During the phase of collecting lesson plans, there were five Taiwanese STEM lesson plans collected, while 14 Indonesian STEM lesson plans were obtained. Lastly, eight teachers from Indonesia participated in the teacher interview process, while four Taiwanese teachers expressed their willingness to be interviewed. Figure 6 provides an overview of the number of participants involved in each type of data collection.

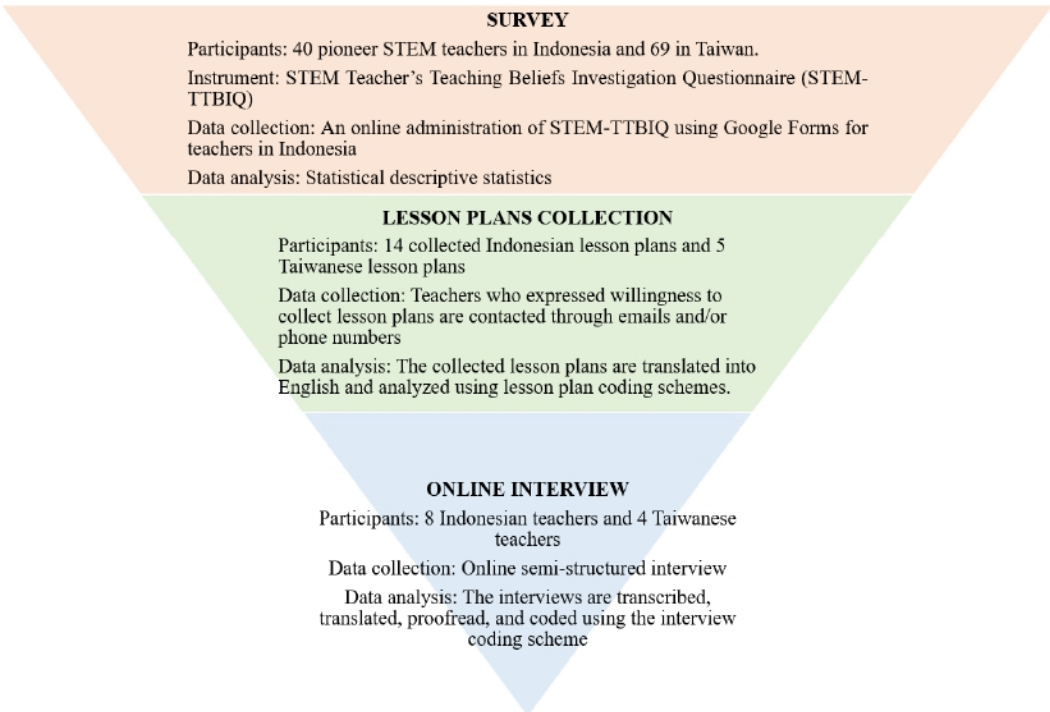

Figure 6. Data pool of study

Although the sample sizes differ, this comparative study on STEM teachers in Indonesia and Taiwan, including 69 Taiwanese teachers and 40 teachers from Indonesia, is justified for several reasons. Firstly, having a representative sample that captures the characteristics and qualities of STEM teachers in each context is crucial. The selection process ensures equity by including participants who meet the established criteria, regardless of the number difference. Secondly, practical considerations such as the availability and accessibility of eligible participants in each country influence the sample sizes. The study aims to gather meaningful insights from both groups by working within the available resources. Thirdly, qualitative data collection methods, such as interviews or observations, can provide deeper insights into the similarities and differences between the teachers from Taiwan and Indonesia, compensating for any potential limitations arising from the variation in sample sizes. Thus, including 69 Taiwanese and 40 Indonesian teachers who meet the established criteria ensures practicality, representative sampling, and the ability to conduct valuable comparative analysis in this study.

In the quantitative phase of the study, a sample of 40 Indonesian teachers and 69 Taiwanese teachers was included. While determining sample size in mixed-method studies does not adhere to fixed rules, considerations such as available resources, time constraints, and practical feasibility are important. Although this sample size may not represent the entire Indonesian and Taiwanese STEM teachers' population statistically, it can still provide valuable insights and preliminary conclusions within the study's specific context.

Acknowledging the limitations of the sample size and its potential impact on the generalizability of the findings is essential. By transparently discussing the rationale behind the chosen sample size and the practical considerations involved, a thoughtful approach to research design is demonstrated. It is worth emphasizing that the quantitative phase serves as a valuable complement to the qualitative findings, offering a broader perspective and enabling the identification of potential trends and relationships within a larger sample.

In mixed-method studies, the sequence of research phases may vary, and conducting the quantitative phase first is a valid approach. By incorporating quantitative and qualitative methods, the study contributes to a comprehensive understanding of integrated STEM education in the Indonesian and Taiwanese contexts. Moreover, future research with larger sample sizes or specific sampling techniques may be necessary to validate further and generalize the findings obtained in the study.

Regarding the classification of participants as Indonesians and Taiwanese in this study, it is important to address the appropriateness of this categorization. Firstly, the participant selection criteria were meticulously designed to ensure the inclusion of experienced STEM teachers with relevant teaching backgrounds in STEM education. The study aimed to gather valuable insights into instructional practices and STEM teaching beliefs by prioritizing individuals with first-hand experiences in teaching STEM. This selection process focused on expertise and pedagogical approaches rather than specific school or regional affiliations.

While the study did not explicitly examine the diversity of subject matter taught by the participants, it is crucial to acknowledge that including STEM teachers from different disciplines within the Indonesian and Taiwanese contexts could contribute to a broader understanding of integrated STEM education. Although the specific subject areas taught were not the primary focus, teachers specializing in various STEM disciplines imply a range of instructional approaches, content knowledge, and perspectives within the integrated STEM education framework. This diversity of backgrounds and expertise may have indirectly influenced the findings, offering a multifaceted view of the challenges and opportunities associated with integrated STEM education across different subject areas. Thus, while not explicitly examined, the diversity of subject matter taught by the participants potentially enhances the richness and applicability of the study's findings.

Moreover, it is important to acknowledge the study’s limitations, particularly the absence of specific school or regional information. However, identifying participants as Indonesians and Taiwanese is justified based on the recruitment criteria, the focus on teaching beliefs within the STEM learning context, the diverse range of participants, and the ethical considerations employed. The participants' willingness to provide consent and actively engage in the study indicates their commitment to advancing STEM education practices. Their invaluable insights contribute to the overall understanding of the research topic.

In qualitative research, identifying participants as Indonesian and Taiwanese STEM teachers aligns with obtaining a deep exploration and understanding of the research phenomena. Qualitative studies often employ purposive or convenience sampling to capture diverse perspectives and rich data within a specific research context. By recruiting participants from Indonesia and Taiwan, this research seeks to shed light on the experiences and perspectives of STEM teachers within their respective educational systems.

Furthermore, considering participants as Indonesians and Taiwanese provides contextual relevance, allowing for a deeper exploration of the unique cultural and educational factors that may influence STEM teaching practices in these countries. The distinct characteristics, policies, and cultural contexts of the Indonesian and Taiwanese education systems shape the implementation of STEM education. Focusing on STEM teachers from these contexts uncovers the nuanced ways in which integrated STEM education is understood, embraced, and enacted within their educational landscapes. This contextual understanding is crucial for developing targeted interventions and strategies to enhance STEM education in these countries.

Additionally, by delving into the experiences and perspectives of STEM teachers in Indonesia and Taiwan, the study acknowledges the localized nature of educational practices. It recognizes the importance of capturing the specificities of these contexts. The findings offer valuable insights into the challenges, successes, and cultural nuances surrounding integrated STEM education in Indonesia and Taiwan. These insights may not be fully captured in more generalized studies that overlook the unique characteristics of these educational systems.

Furthermore, including participants from Indonesia and Taiwan facilitates cross-cultural comparisons and the identification of potential similarities and differences in STEM teaching practices. Comparative analyses can yield valuable insights contributing to the broader understanding of integrated STEM education and foster future international collaborations and knowledge sharing.

STEM education goals in this questionnaire consist of 21st-century skills, STEM literacy, and interdisciplinary problem-solving (Han et al., 2021). The teacher self-efficacy (Enochs & Riggs, 1990) and commitment (Meyer et al., 1993) subscale was revised and adopted from previous research and, in this study, consists of implementers, designers, and disseminators. The participants were asked to respond to STEM education goals, teacher self-efficacy, and commitment on a 5-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). The background information of the STEM teachers, such as their demographics, total years of teaching, and STEM teaching experience, was asked in the posterior parts. STEM-TTBIQ is attached in Appendix A.

The STEM-TTBIQ questionnaire, initially developed in Chinese Traditional, was provided with an English version by the developers (attached in Appendix B) to facilitate its use in this study. Before translating the STEM-TTBIQ questionnaire for this research, the developers granted permission for its utilization. Subsequently, proficient researchers proficient in Indonesian and English languages translated from English to Indonesian, ensuring accuracy and cultural appropriateness. The translated questionnaire was then administered to Indonesian teachers. To assess the validity and reliability of the Taiwanese and Indonesian versions of the STEM-TTBIQ, researchers conducted validity and reliability tests utilizing Rasch analysis. This comprehensive approach ensured the appropriateness and consistency of the questionnaire across both Taiwanese and Indonesian contexts. The Rasch Analysis Model is a widely recognized approach for evaluating the psychometric properties of instruments, particularly in educational and social sciences research (Lah & Tasir, 2018). The researchers used the Rasch Analysis Model as a psychometric test to validate and assess the instrument's reliability (Alnahdi et al., 2021; Dari et al., 2023). This model offers precise and consistent measures, and it has been employed in previous studies to assess the validity and reliability of questionnaires (Gao, 2020; Scoulas et al., 2021).

The Rasch Analysis Model (RAM) is a powerful tool for analyzing data, allowing researchers to determine item and person reliability (Dari et al., 2023; Lah & Tasir, 2018; Scoulas et al., 2021). Item reliability assesses the quality of the items used in a questionnaire, while person reliability provides information about the consistency of participants' responses. In addition, Cronbach's alpha is calculated to assess internal consistency. To determine person reliability, the criteria used are outfit mean square (MNSQ) and outfit Z-standard (ZSTD). The acceptable values for MNSQ and ZTSD are between 0.05 and 1.5 and -2.00 and 2.00, respectively (Alnahdi et al., 2021). The summary statistics of the Indonesian version of the STEM-TTBIQ were obtained and presented in Table 1.

Table 1. Rasch Model Analysis Summary Statistics of STEM-TTBIQ Indonesian Version

Measure		Infit		Outfit		Reliability	
		MNSQ	ZSTD	MNSQ	ZSTD	Real	Model
Person						.93	.93
Mean	4.49	1.06	.15	.99	-.17		
SD	1.79	.36	1.97	.42	1.84		
Item						.86	.87
Mean	.00	.97	-.16	.99	.07		
SD	1.05	.38	1.32	.46	1.29		
Cronbach’s alpha person raw score "test" reliability = .98							

The overall reliability value, shown by Cronbach’s alpha of 0.98, indicates that the Indonesian version of STEM-TTBIQ is reliable (Taber, 2018). Table 2 also shows that the person reliability values for the Indonesian version of the STEM-TTBIQ fall into the high category at 0.93. Additionally, the item reliability values of 0.86 and 0.87 indicate good reliability. The INFIT MNSQ and OUTFIT MNSQ values indicate the respondents' suitability for this research. The values for the teachers' outfits MNSQ and ZSTD are 1.06 and 0.99, respectively, close to the ideal value of 1.00. These values show that the teachers are suitable respondents for this research. The criteria for item suitability are shown by outfit mean square (MNSQ) and outfit (ZSTD). The acceptable value of outfit MNSQ is between 0.5 < MNSQ < 1.5, and the acceptable value of outfit ZSTD is between -2.0 < ZSTD <+ 2.0. The outfit MNSQ and ZSTD values of 0.99 and 0.07, respectively, indicate that the items suit this research. After conducting the validity and reliability test on the Indonesian version of the STEM-TTBIQ, the researchers also analyzed the data from the Traditional Chinese version of the STEM-TTBIQ, administered to Taiwanese teachers. Table 2 presents the summary statistics of the Traditional Chinese version of the STEM-TTBIQ.

Table 2. Rasch Model Analysis Summary Statistics of STEM-TTBIQ Chinese (Traditional) Version

Measure		Infit		Outfit		Reliability	
		MNSQ	ZSTD	MNSQ	ZSTD	Real	Model
Person						.97	.98
Mean	1.88	.99	-.63	1.06	-.52		
SD	1.41	.70	3.05	.79	3.14		
Item						.96	.97
Mean	.00	.98	-.24	1.06	.04		
SD	1.10	.33	1.77	.42	1.98		
Cronbach's alpha person raw score "test" reliability = .98							

The overall reliability value, assessed by Cronbach's alpha, was found to be 0.98, indicating the instrument's reliability. Table 2 presents person reliability values of 0.97 and 0.98 for the Traditional Chinese version of STEM-TTBIQ, which fall into the high category, indicating that the respondents are consistent in their answers. Moreover, the item reliability values of 0.96 and 0.97 indicate good item reliability. The INFIT MNSQ and OUTFIT MNSQ values further reveal that the recruited Taiwanese teachers are suitable respondents for this research. The values for outfit MNSQ and ZSTD, which are 0.99 and 1.06, respectively, are close to the ideal value of 1.00, showing that the respondents are well-suited for this research. In addition, the criteria for item suitability are evaluated using outfit MNSQ and ZSTD values. The acceptable value for outfit MNSQ is between 0.5 and 1.5, and the acceptable value for outfit ZSTD is between -2.0 and +2.0. The outfit MNSQ and ZSTD values of 1.06 and 0.04, respectively, indicate that the items in the Traditional Chinese version are appropriate for use in this research for Taiwanese teachers.

A unidimensional test ensures that the instrument accurately measures STEM Teaching Beliefs. In this model analysis, determining the validity of the test involves evaluating the instrument's unidimensionality. This is a crucial measurement for assessing the construct validity using Residual Principal Component Analysis (PCA). The test involves determining the range of the raw variance explained by measures and unexplained variance in the 1st contrast. A raw variance explained by measures of >40% and unexplained variance in the 1st contrast of <15% are considered acceptable (Van Zile-Tamsen, 2017).

Table 3. Unidimensionality of STEM-TTBIQ

	Indonesian STEM-TTBIQ	Chinese Traditional STEM-TTBIQ
Raw variance explained by measures	50.9%	54.1%
Unexplained variance in 1st contrast	8.2%	7.7%

Table 3 shows that the raw variance explained by measures is 50.9% for Indonesian STEM-TTBIQ and 54.1% for Traditional Chinese STEM-TTBIQ. The unexplained variance in the 1st contrast is 8.2% for Indonesian STEM-TTBIQ and 7.7% for Traditional Chinese STEM-TTBIQ. Therefore, this analysis reveals that the raw variance explained by measures is good, with a value greater than 40%, while the unexplained variance in the 1st contrast is also acceptable, with a value less than 15%. In conclusion, the items used in the STEM-TTBIQ accurately measure STEM teaching beliefs.

Based on the Rasch Model Analysis results, the Chinese Traditional and Indonesian versions of STEM-TTBIQ have good psychometric properties and are a valid measure of STEM teaching beliefs among Taiwanese and Indonesian teachers. The high-reliability values indicate that the instrument is consistent and dependable in measuring the construct of interest. Additionally, the person reliability values suggest that the respondents are consistent in their responses, while the item reliability values suggest that the items measure what they are intended to measure. The outfit MNSQ and ZSTD values further suggest that the recruited Taiwanese and Indonesian teachers are suitable respondents for this research, and the items are appropriate for use in this context. In the data collection stage of this study, STEM-TTBIQ was administered using Google Forms for Indonesian teachers. The results were then exported into Microsoft Excel. IBM SPSS software was used to analyze the data from the questionnaire to obtain descriptive statistics and t-test analysis.

2. Lesson Plans Analysis Protocol

In this study, the STEM lesson plans used for analysis are the written instructions for teachers to implement STEM teaching in their classrooms. Lesson plans serve as written instructions for teachers, outlining the necessary components for implementing STEM teaching in their classrooms (Hite, 2020). They provide details on core competencies, learning goals, teaching points, instructional activities, and assessment methods (Ndihokubwayo et al., 2022).

Analyzing lesson plans is crucial because it reveals how teachers conceptualize and approach STEM teaching beyond cognitive and contextual considerations. It allows researchers to explore the alignment between learning goals and the strategies employed to achieve them. Understanding teachers' beliefs through their lesson plans yields valuable insights into their instructional practices and the factors that shape their teaching designs.

The analysis encompassed five Taiwanese and eight Indonesian lesson plans. Informed consent was obtained from all participants who contributed to collecting their STEM lesson plans. The research team, possessing proficiency in both English and Indonesian and Traditional Chinese languages, translated and proofread the Taiwanese and Indonesian lesson plans. This meticulous process aimed to ensure the accuracy of the English translations to the original lesson plans. During the analysis phase, two coders participated, utilizing English as the working language. Among the coders, one was a native Indonesian speaker, while the other was a native Taiwanese speaker. The collected lesson plans were coded using a designated coding scheme.

The coding scheme in this study was developed based on established theoretical frameworks in qualitative research, ensuring a systematic and rigorous approach to data analysis. By aligning with these theories, the study generated meaningful insights and built upon existing knowledge in the field. The coding scheme focused on the STEM-TTBIQ dimensions of teaching approaches, STEM educational goals, and STEM instructional designs. STEM educational goals included 21st-century skills, STEM literacy, interdisciplinary problem-solving, engagement, and career choice. STEM integration levels encompassed multidisciplinary, interdisciplinary, and transdisciplinary approaches. STEM instructional designs covered inquiry-, problem-, project-, and design-based learning.

STEM instructional designs are the specific strategies and techniques used to teach and facilitate learning in STEM subjects. Effective STEM instructional designs should align with the chosen educational goals and integration level. For example, suppose the educational goal is to develop problem-solving skills, and the integration level is transdisciplinary. In that case, the STEM instructional design should include real-world problems that require students to draw on knowledge and skills from multiple STEM disciplines to find a solution. In short, STEM instructional designs should be carefully planned and designed to align with STEM educational goals and the chosen integration level to ensure that students receive a meaningful and engaging STEM education.

This study analyzed the lesson plans' dimensions of teaching approaches and STEM educational goals. These constructs were easily identifiable within the written artifacts. It is important to recognize that lesson plans alone may not accurately reflect a teacher's self-efficacy or commitment. Therefore, we decided to limit the coding scheme to observable constructs within the lesson plans. This approach did not compromise the analysis's integrity but allowed us to develop a comprehensive and relevant coding scheme aligned with our research questions. By adopting this top-down approach and focusing on the most relevant constructs, we enhanced the validity and reliability of our analysis.

The top-down approach was used to analyze the lesson plans, meaning the coding scheme was developed first and then applied to the lesson plans. This approach involves identifying pre-established codes or categories according to STEM-TTBIQ dimensions and using them to guide the coding process (Rauss & Pourtois, 2013). The predetermined categories and themes identified facilitated the coding process, promoting replicability and comparability of findings (Casimir et al., 2022; Rauss & Pourtois, 2013).

A comprehensive literature review was conducted on STEM Educational Goals, Integration Levels (English, 2016; Margot & Kettler, 2019), and Instructional Designs commonly used in STEM education. STEM Educational Goals were further classified into four sub-dimensions, namely 21st-century skills (17 codes) (Kareem et al., 2022), STEM literacy (3 codes) (Jackson et al., 2021), interdisciplinary problem-solving (9 codes) (Melton et al., 2022), engagement, and career choice (8 codes) (De Loof et al., 2022; Murphy et al., 2019). Each sub-dimension was assigned a set of codes. STEM Integration Level was assigned three codes: multidisciplinary, interdisciplinary, and transdisciplinary (English, 2016; Margot & Kettler, 2019). STEM Instructional Design appointed four types of learning: Inquiry-based learning (Forbes et al., 2020; García-Carmona, 2020; Liu et al., 2021; van Uum et al., 2016); problem-based learning (Dolmans et al., 2005; Kim et al., 2018; Rehmat & Hartley, 2020; Smith et al., 2022), Project-based learning (Beier et al., 2019; Capraro et al., 2013b; Craig & Marshall, 2019; Lu et al., 2022; Tsinajinie et al., 2021), and Design-based learning (Azizan & Abu Shamsi, 2022; van Diggelen et al., 2021; Zhang et al., 2022). The coding scheme of the lesson plan is attached in Appendix C.

The STEM educational goals were coded throughout the entire lesson plan. They were predominantly found in the lesson objectives section, which existed in Indonesian and Taiwanese lesson plans. The Integration Level was identified from the areas of the lesson plans where teachers explicitly stated which types of STEM disciplines were being incorporated. Alternatively, Integration Level was determined by looking for indications of the codes throughout the teaching and learning activity descriptions. Regarding instructional design, the learning activities were scrutinized to identify the codes from four types of instructional designs.

In this study, intercoder reliability (ICR) was utilized to ensure the reliability of the coding scheme used to analyze Indonesian and Taiwanese STEM lesson plans. ICR was used to assess the level of agreement between different coders regarding how they categorize and code the same data (Cheung & Tai, 2021; McDonald et al., 2019). Assessing the intercoder reliability (ICR) impacted the qualitative analysis's quality, transparency, and reception. The coding process's systematicity, communicability, and transparency were improved as multiple coders ensured consistency and accuracy through ICR assessment. It also facilitated reflexivity and dialogue within the research team as the coders could discuss and resolve any discrepancies in their coding. The ICR assessment also provided evidence of the reliability and consistency of the coding process, which helped to convince diverse audiences of the trustworthiness of the analysis (O'Connor & Joffe, 2020).

Standard procedures were followed to ensure the accuracy of the coding process (Cheung & Tai, 2021; McDonald et al., 2019; O'Connor & Joffe, 2020). Two coders underwent training to familiarize themselves with the coding scheme and the code assignment criteria. Moreover, the coders worked independently and later met to discuss discrepancies and reach a consensus. The coding results were compiled, and the percentages of those codes emerging from the analysis of the lesson plans from both countries were calculated. The procedure for coding the lesson plans involved a team of trained researchers reviewing each lesson plan and applying the coding scheme to identify which categories and subcategories were present. The two coders discussed discrepancies in their coding of the lesson plan to come to a consensus. ICR was found to be 0.87, which is considered acceptable. This indicates that the coding scheme used in this study was reliable and consistent across multiple coders.

3. Interview Protocols

The interview is intended to probe teachers' teaching beliefs according to their answers to the questionnaire and lesson plans. The interview questions were designed based on these dimensions to elicit information on various aspects of STEM education. Table 4 includes nine questions covering different areas related to STEM education, such as defining STEM, factors that interest teachers in implementing STEM learning, and using teaching strategies and media in STEM lessons.

Table 4. Teacher Interview Questions

Interview Questions
1. How do you define STEM?
2. What factors make you interested in STEM education and implementing STEM learning?
3. How does the current curriculum guideline facilitate you with the opportunity to implement STEM learning?
4. What do you think about the students' abilities that can be improved by STEM learning? How do you design your lesson plan to achieve the objective?
5. According to your lesson plan, please describe the teaching approach you used when implementing STEM learning and why you chose that specific approach.
6. According to your lesson plan, what specific strategies and teaching media have you chosen for your STEM lessons? Why?
7. How do you assess students' learning performance and outcomes in STEM lessons?
8. What factors do you think are essential to be a professional in STEM education?
9. How do you see yourself improving your capabilities in delivering STEM learning since you have practiced STEM in your teachings?

The interview questions were carefully designed to explore various dimensions of STEM education and provide a comprehensive understanding of participants' perspectives and experiences. The questions were designed to elicit valuable insights into how participants define STEM, their motivations for incorporating STEM learning, and their perceptions of the role of the curriculum guidelines in facilitating STEM implementation. Additionally, the questions delved into participants' perspectives on students' abilities that can be enhanced through STEM learning and how they design their lesson plans to achieve these objectives. The inquiry also extended to participants' chosen teaching approaches, strategies, and teaching media for implementing STEM lessons and their methods of assessing student learning outcomes. Furthermore, the questions probed participants' views on the essential qualities of professionals in STEM education and their plans for personal growth and improvement in delivering STEM learning.

In total, eight Indonesian teachers and four Taiwanese teachers were interviewed. Teachers were recruited based on their willingness to participate in collecting lesson plans and being interviewed. Informed consent was obtained from all participating teachers. The teacher interviews were recorded and then transcribed. Proofreading was done both before and after transcription. The transcripts were then translated into English and proofread again. After the transcripts were ready in English, the coding was conducted following standard procedures for the analysis. The interview coding scheme is attached in Appendix D.

The coding scheme was developed using pre-determined codes from literature review and content analysis. The codes for each question were as follows: Teachers' Understanding of STEM (7 codes) (Kelley & Knowles, 2016), Teachers' Motivation to Implement STEM (9 codes) (Conradty & Bogner, 2020; Jungert et al., 2020; Vermote et al., 2020), The Curriculum Support of STEM Implementation (6 codes), Potential Students' Abilities to be Improved by STEM Learning (33 codes), Teachers' STEM Instructional Design (10 codes), Teaching Media for STEM (4 codes) (Arici et al., 2019; Li et al., 2019; Tawbush et al., 2020; Van Den Beemt et al., 2020), The Assessment Used in STEM Learning (3 codes) (Capraro et al., 2013a; Karakaya & Yılmaz, 2022; Salvo-Garrido et al., 2022), Teachers' Perceptions of the Characteristics of STEM Professionals (14 codes), and Teachers' Professional Development Plan (12 codes).

The intercoder reliability (ICR) result of 0.87 for the interview coding scheme in this study demonstrates a high level of agreement between the coders. Conducting ICR in interview analysis enhances the research process's reliability, transparency, and rigor. It identifies areas for improvement in the coding scheme, ensuring comprehensive and accurate data analysis. The ICR result provides evidence of the coding process's consistency and accuracy, bolstering the trustworthiness of the analysis. It also fosters reflexivity and facilitates dialogue within the research team to resolve coding discrepancies effectively. Following the coding stage, emerging themes were identified based on the percentages of codes derived from teacher interviews in both countries, and the results were compiled, organized, and interpreted.

The process of data triangulation in this study involves the use of three different data sources to investigate STEM teachers' teaching beliefs. By employing a triangulation approach, researchers comprehensively understand the phenomena under study by considering multiple perspectives. This study combines quantitative and qualitative methods to collect data, allowing for a more in-depth exploration of the research topic. This approach enhances the robustness and validity of the findings, as integrating multiple data sources provides a more nuanced and holistic understanding of STEM teachers' teaching beliefs. The integration of all data collected is shown in Figure 7 as follows.

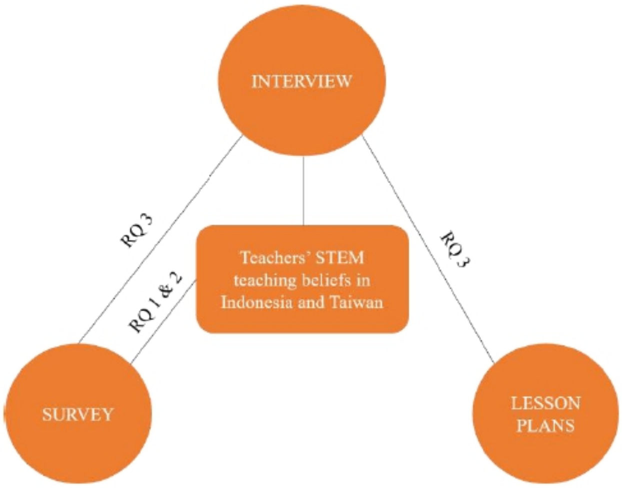

Figure 7. Triangulation of data