INTERNATIONAL JOURNAL OF ACADEMIC RESEARCH IN PROGRESSIVE EDUCATION & DEVELOPMENT
OPEN ACCESS JOURNAL
Vol 14, Issue 3, (2025) E-ISSN: 2226-6348

A Conceptual Framework for Evaluating the Implementation of the Integrated Science Curriculum in Zhejiang Province, China Using the CIPP Model

Zeng Xiangfei[1], Nor Hasnida Che Md Ghazali[2]*, Yao Yao[3], Huang Dongyuan[4]
[1,2]Faculty of Human Development, Universiti Pendidikan Sultan Idris, Malaysia, [3]Senior High School Division, Hainan Micro-city Future School, China, [4]Junior High School Division, Hainan Micro-city Future School, China
Email: 2421645693@qq.com[1], 1016126810@qq.com[3], 2199162749@qq.com[4]
*[2]Corresponding Author Email: hasnida@fpm.upsi.edu.my

To Link this Article: http://dx.doi.org/10.6007/IJARPED/v14-i3/25566 DOI:10.6007/IJARPED/v14-i3/25566

Published Online: 08 July 2025

Abstract
This conceptual paper proposes an evaluation framework to assess the implementation of the Integrated Science Curriculum in lower secondary schools in Zhejiang Province, China. In response to national education reforms aimed at enhancing scientific literacy, the study adopts the Context, Input, Process, and Product (CIPP) model to develop a structured and multidimensional evaluation framework. The framework is developed based on the Compulsory Education Science Curriculum Standards (2022 Edition) and existing evaluation studies related to integrated science curricula. It identifies key indicators reflecting curriculum relevance, resource provision, instructional practices, and learning outcomes. The framework is theoretically supported by constructivism, neuroscience, and complex systems theory. Methodologically, this study adopts a conceptual approach and does not involve empirical data collection. It outlines a CIPP-based structure and proposes a questionnaire-based method for future empirical validation. The framework contributes both theoretically and practically to advancing integrated science curriculum evaluation and provides implications for science educators, policymakers, and school administrators.
Keywords: Integrated Science Curriculum, CIPP Model, Curriculum Evaluation, Scientific Literacy, Conceptual Framework

Introduction
In the 21st century, scientific literacy has emerged as a critical competency for personal development, national competitiveness, and social sustainability. The Outline of the National Action Plan for Scientific Literacy (2021–2035) emphasizes that cultivating scientifically literate citizens is essential for building an innovative country (State Council of China, 2021). Scientific literacy includes not only understanding scientific concepts but also the ability to

inquire, reason, and make evidence-based decisions in real-world contexts (Bybee, 1997; Ministry of Education of China [MOE], 2022). To meet this educational demand, China has continuously advanced science curriculum reform at the compulsory education level. The latest Compulsory Education Science Curriculum Standards (2022 Edition) propose the integrated development of students' scientific literacy across scientific knowledge, scientific thinking, scientific inquiry and practice, and scientific attitude and responsibility (MOE, 2022). Promoting integrated science curriculum reform is a key strategy to enhance scientific literacy (Åström, 2008). In China, an integrated science curriculum is often defined as a comprehensive interdisciplinary curriculum that includes elements of physics, chemistry, biology, and natural geography. As one of the earliest provinces to implement such reforms, Zhejiang has offered integrated science curriculum at the lower secondary level since 1988 and remains a national exemplar in breadth and depth of implementation.

Despite ongoing reforms, the implementation of the integrated science curriculum in Zhejiang Province continues to face several enduring challenges. Many science teachers, shaped by long-standing disciplinary training systems, find it difficult to adapt to interdisciplinary teaching approaches required by the integrated curriculum (Jia, 2013). Early development of curriculum materials also lacked coherence and integration, leading to fragmented instructional content (Fei, 2012; Pan, 2004). Moreover, significant disparities exist between urban and rural schools in terms of teacher qualifications, resource allocation, and instructional practices, leading to inconsistent science curriculum delivery (Miao et al., 2024; Shen, 2024; Zhang, 2022). Most critically, there is still a lack of systematic evaluation mechanisms tailored to the integrated science curriculum. Existing assessment frameworks remain grounded in subject-based traditions and fail to comprehensively measure integrated science curriculum effectiveness across planning, resource input, instructional process, and student outcomes (Wang, 2024).

This study aims to evaluate the implementation of the Integrated Science Curriculum in lower secondary schools in Zhejiang Province based on the Context, Input, Process, and Product (CIPP) model as its conceptual evaluation framework. The framework is intended to identify gaps between policy intentions and classroom realities, clarify multidimensional influencing factors, and provide a theoretical basis for developing future empirical tools and strategies for curriculum improvement.

Literature Review

Rationale for Choosing the CIPP Model

Curriculum evaluation refers to the systematic collection and analysis of information related to educational curricula to determine their effectiveness, value, and impact (Glatthorn et al., 2018). Unlike student assessment, which focuses on individual academic performance, curriculum evaluation offers a holistic examination of educational objectives, implementation processes, and learning outcomes (Vaidya, 2014). In science education, curriculum evaluation is particularly critical for ensuring alignment with national standards, responding to student needs, and enhancing scientific literacy.

Over the past century, curriculum evaluation has progressed from measurement-focused approaches to more diversified models. In the early 20th century, the Measurement Era emphasized standardized testing and quantitative outcomes, represented by Edward

Thorndike (1904). This shifted in the 1930s with Tyler's Objectives Model, which aligned evaluation with specific learning goals, though it often overlooked contextual factors and learner diversity (Tyler, 1934). Later models, such as Stake's Responsive Evaluation and Scriven's Goal-Free Evaluation, addressed real-world complexity and unintended outcomes. However, Stake's model lacked a clear structure, limiting its replicability across contexts, while Scriven's model, though innovative, provided little procedural guidance and was less useful for informing policy (Scriven, 1991; Stake, 1976). More recent constructivist approaches, like Guba and Lincoln's Fourth Generation Evaluation, emphasize negotiation and stakeholder participation, but their qualitative nature and emphasis on consensus-building may reduce applicability in large-scale evaluations (Guba & Lincoln, 1989).

In contrast, the CIPP Model—comprising Context, Input, Process, and Product—combines a systematic structure with a practical, decision-oriented function. It supports both formative and summative evaluation and facilitates continuous feedback throughout educational reform (Stufflebeam, 2003). Despite its operational complexity and the need for a high level of methodological sophistication, the model offers a comprehensive approach to evaluating and improving educational programs. Although empirical applications of the CIPP model in China's integrated science curriculum remain limited, it has been widely adopted internationally to evaluate science and STEM-related curricula (Aziz et al., 2018; Masfuah et al., 2022). In the domestic context, scholars such as Zhuo (2017) have applied similar multidimensional frameworks to evaluate interdisciplinary instruction.

Applications of the CIPP Model in Science Education

Several studies have examined the application of the CIPP model in evaluating integrated science curricula. The context dimension focuses on the alignment between the curriculum and its external environment, including national standards, student needs, and curricular objectives. Compliance with national science curriculum standards is a key determinant of a curriculum's legitimacy and scientific validity (Arjaya & Suma, 2023; MOE, 2022). Furthermore, when the curriculum is designed in response to students' actual needs, it not only reduces instructional inefficiencies but also enhances student engagement and interest (Liang, 2024; Ling et al., 2017). In addition, clearly defined and attainable curricular objectives have been shown to improve teacher clarity and instructional effectiveness (Liu & Li, 2024; San Diego STEM Ecosystem, 2014; Zhang, 2019).

The input dimension assesses the foundational resource conditions necessary for effective curriculum implementation. Key components identified in the literature include the scientific and pedagogical quality of teaching content and instructional materials (Aziz et al., 2018; Zhang, 2019), the adequacy of laboratory infrastructure and school environments (Arjaya & Suma, 2023; Jibril & Bagceci, 2024; Liang, 2024), and the qualifications and interdisciplinary teaching competencies of science teachers (Bashri et al., 2020; Masfuah et al., 2022). These factors together determine the feasibility and support level for implementing the integrated science curriculum.

The process dimension assesses how the curriculum unfolds in classroom instruction. Studies emphasize that classroom management—especially in terms of lesson organization and time control—directly affects implementation quality (Rooholamini, 2017). Pedagogical adaptability, such as the integration of lecture-based, inquiry-based, practice-based, or blended approaches, is essential to meet students' learning needs (Arjaya & Suma, 2023; Kereeditse, 2021). Student engagement, including behavioral, emotional, and cognitive involvement, serves as a key indicator of instructional effectiveness ((Liu & Li, 2024; Zhang, 2019). Reflective practices by both teachers and students foster self-regulation and pedagogical improvement (Jibril & Bagceci, 2024). Feedback—both received by students and used by science teachers to refine instruction—further supports continuous curriculum improvement (Bashri et al., 2020).

The product dimension evaluates the degree to which the curriculum achieves its educational objectives, particularly in terms of students' scientific literacy. According to the 2022 edition of China's Compulsory Education Science Curriculum Standards, scientific literacy encompasses four core dimensions, which are widely supported in empirical studies and provide a reliable basis for outcome evaluation (Liang, 2024; Zhuo, 2017).

Conceptual Framework

A conceptual framework refers to a model or schematic of the relationships between core concepts and variables in a study (Jaakkola, 2020). This study constructs a conceptual framework based on prior research on science curricula and the CIPP model to evaluate the implementation of the integrated science curriculum in Zhejiang Province. Each dimension is represented by a set of grounded indicators. The context dimension includes three factors—science curriculum standards, student needs, and science curriculum objectives—which reflect the foundation and background of curriculum design. The input dimension consists of curriculum content, learning environment, teaching resources, and science teachers, highlighting the material and human conditions required for curriculum implementation. The process dimension incorporates five indicators: management of teaching activities, adaptive teaching methods, student engagement, reflection, and feedback, which comprehensively assess instructional quality and classroom dynamics. Finally, the product dimension includes scientific understanding, scientific thinking, scientific inquiry and practice, and scientific attitude and responsibility, providing a structured way to measure the outcomes of science education. By combining multiple indicators across the four dimensions, this framework enables a comprehensive and balanced evaluation of curriculum implementation.

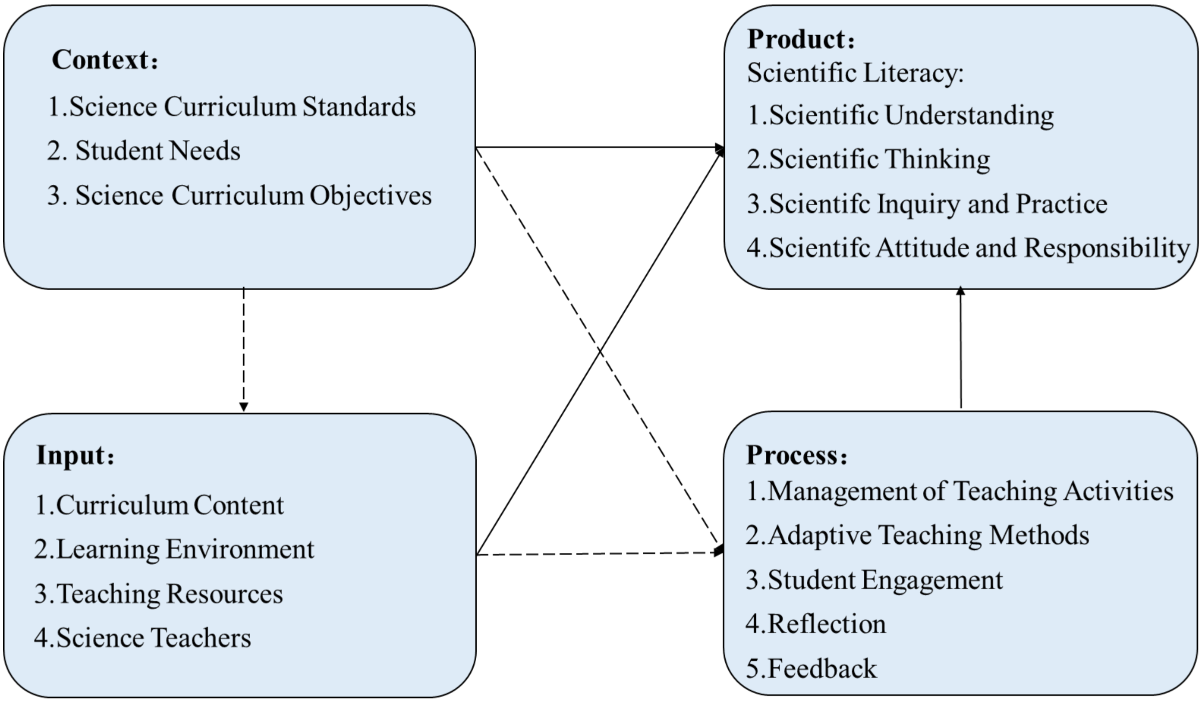

Figure 1: Proposed Conceptual Framework

Theoretical Foundation

Constructivism provides a foundational perspective for this study, emphasizing that learners actively construct knowledge through engagement with their environment and by integrating new experiences with existing cognitive structures (Piaget, 1997; Vygotsky, 1978). Piaget's theory of cognitive development highlights the importance of developmental readiness, while Vygotsky's sociocultural theory underscores the role of social interaction in learning. This view supports curriculum designs responsive to students' developmental stages, cultural backgrounds, and inquiry capacities. In curriculum evaluation, constructivist principles have informed the development of student-centered, participatory approaches aligned with real-life contexts. For instance, G. Zhang et al. (2011) and Frye and Hemmer (2012) demonstrated how reflective practice and active involvement—central to constructivist thinking—can enrich educational program evaluation. These studies support integrating constructivist ideas into comprehensive evaluation frameworks such as the CIPP model (Stufflebeam, 2003).

Neuroscience explores how the brain supports learning through mechanisms such as attention, memory, and neuroplasticity—the brain's ability to adapt through experience (Voss et al., 2017). Educational neuroscience applies these insights to instructional design, emphasizing the importance of emotionally engaging and cognitively appropriate learning environments aligned with brain development (Villa, 2018). These insights contribute to curriculum evaluation by offering a scientific foundation for understanding how learners process information, engage with instruction, and develop skills. For example, Wang et al. (2023) proposed a neural model of classroom information processing, revealing how students acquire and retain knowledge, informing teaching and evaluation strategies. Similarly, Howes et al. (2013) highlighted the interplay between biological and environmental factors in shaping learning outcomes, suggesting that curriculum design and evaluation should consider individual neurodevelopmental differences. As such, neuroscience provides strong theoretical support for the conceptual framework, particularly in understanding the mechanisms underlying effective curriculum implementation and evaluation.

Complex systems theory conceptualizes curriculum implementation as a dynamic and emergent process arising from the complex interactions among students, teachers, curriculum content, and the learning environment (Davis & Sumara, 2014; Mitchell, 2009). Unlike traditional linear models, this perspective emphasizes the importance of nonlinearity, adaptability, and feedback loops within educational systems (Boulton et al., 2015). Byrne and Callaghan (2022) highlighted the crucial role of systemic feedback mechanisms in optimizing classroom dynamics, while Davis and Sumara (2014) further emphasized the importance of flexible instructional design in enhancing student engagement. Complex systems theory aligns closely with the core features of the CIPP curriculum evaluation model, thereby providing strong theoretical support for comprehensive and developmental curriculum evaluation.

Operational Definitions

Operational definitions are clear and specific definitions of variables or indicators in terms of perceptible and measurable things, events, phenomena, and methods. In addressing the context dimension, the integrated science curriculum is required to align with national science curriculum standards in terms of its nature, concept, objectives, content, academic quality, and implementation to foster the comprehensive development of scientific literacy (MOE, 2022). Student needs are conceptualized as the discrepancy between current and expected learning states (Grant, 2002), encompassing prior knowledge and skills (Y. Wang et al., 2019), cognitive readiness (Eilks & Hofstein, 2017), learning styles (Eilks & Hofstein, 2017), and science-related interests (Penuel et al., 2022). Science curriculum objectives refer to the particular outcomes or expected results that students are expected to achieve through learning in science education (Soysal, 2022). Furthermore, curriculum objectives should be clear (McComas, 2024), structured (McComas, 2024), and assessable to provide effective instructional guidance (Michael et al., 1957).

With regard to the input dimension, curriculum content refers to the comprehensive body of information, skills, values, and attitudes systematically delivered to students throughout the educational process to achieve specific teaching objectives (Glatthorn et al., 2018). It should be logically sequenced (Jin et al., 2019), interdisciplinary in nature (Ashby & Exter, 2019), and rooted in real-world applications (Rios & Stanton, 2011). The learning environment encompasses the external social, psychological, and physical conditions and factors that influence students' learning effectiveness and experience (Closs et al., 2022). This study specifically focuses on evaluating the physical learning environment, including the adequacy of classroom layouts, laboratory facilities, and large-scale experimental equipment (Darman, 2023; Pareek, 2019). Teaching resources comprise all materials and tools that facilitate teaching and learning, including textbooks, teaching aids, digital platforms, multimedia tools, and library resources (Yara & Omondi, 2010). These resources are categorized into digital (Heine et al., 2023), textual (Wanselin et al., 2022), and physical types (Fernandez et al., 2024) in this study. Science teachers are defined as professionals specializing in delivering the Integrated Science Curriculum at the lower secondary school level, typically possessing academic backgrounds in disciplines such as physics, chemistry, biology, and geography. Their evaluation considers their professional backgrounds (Margot & Kettler, 2019), teaching skills

(Spektor-Levy et al., 2008), interdisciplinary integration abilities (Wu et al., 2024), and participation in systematic teacher training programs (Wu et al., 2024).

In the process dimension, the management of teaching activities refers to the systematic planning and coordination of teaching content, instructional methods, time scheduling, classroom management, and resource allocation by teachers (Rakhmonkulov & Usarov, 2019). Effective management ensures that teaching content is organized, learning progresses in an orderly manner, student engagement is stimulated, and assessment and feedback are conducted in a timely fashion (Hagermoser Sanetti et al., 2018; Laveault & Allal, 2016). Adaptive teaching methods involve the selection and application of instructional strategies appropriate to the curriculum content, including lecture-based, inquiry-based, practice-based, and blended approaches (Wieman, 2014). Inquiry- and practice-based methods are particularly emphasized in integrated science curricula and national education standards (Aidoo, 2023; MOE, 2022). Student engagement refers to the behavioral, affective, and cognitive commitment demonstrated during the learning process (Schnitzler et al., 2021). Specifically, behavioral engagement involves active participation and task completion; affective engagement reflects students' emotional responses to learning activities; and cognitive engagement indicates deep thinking and critical analysis. Curriculum reflection denotes the systematic review, analysis, and evaluation of learning content, methods, processes, and outcomes conducted by both students and teachers (Suphasri & Chinokul, 2021). It aims to identify achievements, diagnose challenges, and determine areas for improvement, thereby fostering critical thinking, self-regulation, and instructional enhancement. Curriculum feedback refers to the bidirectional exchange of information between students and teachers concerning learning content, teaching methods, and outcomes, with the aim of optimizing teaching effectiveness and improving learning experiences (Nieminen & Carless, 2023). From the student perspective, feedback evaluates the effectiveness of instruction; from the teacher perspective, it informs curriculum adjustments and pedagogical improvements.

Finally, the product dimension centers on the extent to which students develop core competencies in scientific literacy (MOE, 2022). Scientific understanding refers to the overall understanding of objective things formed on the basis of the understanding of scientific concepts, laws, and principles. Scientific understanding includes both specific concepts in the fields of science, technology, and engineering; and the understanding of the nature of science, the relationship between human beings and nature, and the relationship between science, technology, society, and the environment. In addition, scientific understanding covers their applications in explaining natural phenomena and solving practical problems. Scientific thinking is a way of cognising the essential attributes, internal laws, and interrelationships of objective things from a scientific perspective, which mainly includes model construction, reasoning and argumentation, and innovative thinking. The scientific inquiry and practice mainly refer to the scientific inquiry competence, technical and engineering practice competence, and self-directed learning competence developed in the process of understanding and exploring nature, acquiring scientific knowledge, solving scientific problems, and carrying out technical and engineering practice. Scientific attitude and responsibility encompass the development of a scientific attitude and a sense of social responsibility, both of which gradually form based on an understanding of the nature and principles of science, as well as the interconnections between science, technology, society, and the environment.

Methodology

This study adopts a quantitative research design to construct and validate the evaluation framework. A structured questionnaire will serve as the primary research instrument, adapted from previous studies that have applied the CIPP model to evaluate curricula related to integrated science. The target population comprises full-time lower secondary integrated science teachers across Zhejiang Province. To ensure representativeness and minimise sampling bias, simple random sampling will be employed, allowing every eligible teacher an equal probability of selection. The questionnaire adopts a five-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). Prior to the main survey, a pilot study will be conducted to assess the psychometric properties of the instrument, and revisions will be made based on preliminary findings.

Reliability will be evaluated using Cronbach's alpha, with a threshold of 0.70 indicating acceptable internal consistency reliability for each dimension (Taber, 2018). To ensure validity, a three-tiered validation strategy will be used. Face validity will be established through expert reviews by specialists in science education and curriculum studies. Content validity will be measured using the Content Validity Index (CVI). Construct validity will further be assessed through exploratory factor analysis (EFA) and confirmatory factor analysis (CFA). The suitability of data for factor extraction will be determined using the Kaiser-Meyer-Olkin (KMO) measure and Bartlett's test of sphericity.

Data analysis will be carried out using SPSS 26.0 and AMOS 26.0 software. Descriptive statistics (frequencies, percentages, means, and standard deviations) will be used to summarise demographic characteristics and participants' responses across the CIPP dimensions, including variables such as gender, teaching experience, and school location. EFA will be conducted to explore the underlying factor structure, with items exhibiting factor loadings below 0.60 removed through iterative refinement to ensure unidimensionality and robust construct validity (Jani et al., 2023). CFA will then be used to confirm the measurement model, employing fit indices such as CFI, TLI, RMSEA, and the chi-square/df ratio. Finally, structural equation modeling (SEM) will be applied to examine the interrelationships among the four CIPP dimensions and to test hypothesised paths, with standardised coefficients (β) indicating path strength and unstandardised coefficients (B) used for hypothesis testing.

Conclusion

This conceptual paper proposes an evaluation framework based on the CIPP model to evaluate the implementation of the Integrated Science Curriculum in Zhejiang Province. By reviewing relevant literature and integrating constructivism, neuroscience, and complex systems theory, the study outlines key indicators across the four CIPP dimensions. The framework provides a structured and theoretically grounded basis for future empirical evaluation and supports ongoing efforts to improve curriculum quality and promote scientific literacy through interdisciplinary teaching.

References

Aidoo, B. (2023). Teacher educators' experience adopting problem-based learning in science education. *Education Sciences, 13*(11), Article 11113. doi:10.3390/educsci13111113

Arjaya, I. B. A., & Suma, K. (2023). Problems of biology learning and evaluation analysis at the CIPP model-based higher education level. *Biosfer: Jurnal Pendidikan Biologi, 16*(1), 152–167.

Ashby, I., & Exter, M. (2019). Designing for interdisciplinarity in higher education: Considerations for instructional designers. *TechTrends, 63*(2), 202–208. doi:10.1007/s11528-018-0352-z

Åström, M. (2008). *Defining integrated science education and putting it to test* (Doctoral dissertation). Linköping University, Sweden. Retrieved from ProQuest Dissertations and Theses Global.

Aziz, S., Mahmood, M., & Rehman, Z. (2018). Implementation of CIPP model for quality evaluation at school level: A case study. *Journal of Education and Educational Development, 5*(1), 189–206. doi:10.22555/joeed.v5i1.1553

Bashri, A., Prastiwi, M. S., & Puspitawati, R. P. (2020). CIPP model for curriculum evaluation of biology education. In *Proceedings of the International Joint Conference on Arts and Humanities (IJCAH 2020)* (Vol. 491, pp. 1247–1251). Atlantis Press. doi:10.2991/assehr.k.201201.209

Boulton, J. G., Allen, P. M., & Bowman, C. (2015). *Embracing complexity: Strategic perspectives for an age of turbulence*. Oxford, UK: Oxford University Press.

Bybee, R. W. (1997). *Achieving scientific literacy: From purposes to practices*. ERIC Document Reproduction Service No. ED461491.

Byrne, D., & Callaghan, G. (2022). *Complexity theory and the social sciences: The state of the art*. New York, NY: Routledge.

Closs, L., Mahat, M., & Imms, W. (2022). Learning environments' influence on students' learning experience in an Australian faculty of business and economics. *Learning Environments Research, 25*(1), 271–285. doi:10.1007/s10984-021-09361-2

Darman, D. (2023). The effect of teaching quality and campus facilities on student learning motivation. *The Eastasouth Journal of Learning and Educations, 1*(2), 36–43.

Davis, B., & Sumara, D. (2014). *Complexity and education: Inquiries into learning, teaching, and research*. New York, NY: Routledge.

Eilks, I., & Hofstein, A. (2017). Curriculum development in science education. In K. S. Taber & B. Akpan (Eds.), *Science education: An international course companion* (pp. 169–181). Rotterdam, The Netherlands: SensePublishers. doi:10.1007/978-94-6300-749-8_13

Fei, M. (2012). Strategies for strengthening conceptual integration in the implementation of integrated science curriculum. *Education and Teaching Forum, 8*, 81–82.

Fernandez, C., Hochgreb-Haegele, T., Eloy, A., & Blikstein, P. (2024). Making for science: A framework for the design of physical materials for science learning. *Educational Technology Research and Development, 72*(1), 59–82. doi:10.1007/s11423-023-10340-y

Frye, A. W., & Hemmer, P. A. (2012). Program evaluation models and related theories: AMEE guide no. 67. *Medical Teacher, 34*(5), e288–e299.

Glatthorn, A. A., Boschee, F., Whitehead, B. M., & Boschee, B. F. (2018). *Curriculum leadership: Strategies for development and implementation* (5th ed.). Thousand Oaks, CA: SAGE Publications.

Grant, J. (2002). Learning needs assessment: Assessing the need. *BMJ, 324*(7330), 156–159.

Guba, E. G., & Lincoln, Y. S. (1989). *Fourth generation evaluation*. Newbury Park, CA: Sage Publications.

Jaakkola, E. (2020). Designing conceptual articles: Four approaches. *AMS Review, 10*(1), 18–26. doi:10.1007/s13162-020-00161-0

Jani, W., Razali, F., Ismail, N., & Ismawi, N. (2023). Exploratory factor analysis: Validity and reliability of teacher's knowledge construct instrument. *International Journal of Academic Research in Progressive Education and Development, 12*(1), 944–953.

Hagermoser Sanetti, L. M., Williamson, K. M., Long, A. C., & Kratochwill, T. R. (2018). Increasing in-service teacher implementation of classroom management practices through consultation, implementation planning, and participant modeling. *Journal of Positive Behavior Interventions, 20*(1), 43–59.

Heine, S., Krepf, M., & König, J. (2023). Digital resources as an aspect of teacher professional digital competence: One term, different definitions – a systematic review. *Education and Information Technologies, 28*(4), 3711–3738. doi:10.1007/s10639-022-11321-z

Howes, A., Kaneva, D., Swanson, D., & Williams, J. (2013). *Re-envisioning STEM education: Curriculum, assessment and integrated, interdisciplinary studies*. London, UK: The Royal Society.

Jia, L. (2013). *Research in the installation of integrated science curriculum in junior middle school* (Master's thesis). Shandong Normal University, Jinan, China.

Jibril, S., & Bagceci, B. (2024). Comparative evaluation of the 12th grade chemistry curriculum according to the CIPP evaluation model: Türkiye and Nigeria. *Journal of Research in Education, 2*(1), 1–27.

Jin, H., Mikeska, J. N., Hokayem, H., & Mavronikolas, E. (2019). Toward coherence in curriculum, instruction, and assessment: A review of learning progression literature. *Science Education, 103*(5), 1206–1234.

Kereeditse, M. (2021). Evaluation of strengthening of mathematics and science in secondary education (SMASSE) program in Botswana using the CIPP model. *Mosenodi Journal, 24*(1), 21–41.

Laveault, D., & Allal, L. (2016). *Assessment for learning: Meeting the challenge of implementation* (Vol. 4). New York, NY: Springer.

Liang, Y. Q. (2024). Evaluation practice of kindergarten science curriculum based on the CIPP model. *Educational Science Forum, 16*(6), 31–35.

Ling, L. S., Pang, V., & Lajium, D. (2017, May). *An evaluation of the development and implementation of integrated STEM module using CIPP evaluation model: A proposed study*. Paper presented at the Simposium Psikologi dan Kesihatan Sosial-i (SPKS-i) 2017, Universiti Malaysia Sabah, Kota Kinabalu, Malaysia.

Liu, X. M., & Li, G. (2024). Research on the construction of evaluation system of STEM education implementation in China's primary and secondary schools under the perspective of CIPP. *Research in Teaching, 47*(3), 41–55.

Margot, K. C., & Kettler, T. (2019). Teachers' perception of STEM integration and education: A systematic literature review. *International Journal of STEM Education, 6*(1), 2. doi:10.1186/s40594-018-0151-2

Masfuah, S., Fakhriyah, F., & Hilyana, F. S. (2022). Blended learning based on science literacy in science concept learning. In *AIP Conference Proceedings* (Vol. 2600). Paper presented at the 6th Science Education International Conference (SEIC), Semarang, Indonesia. doi:10.1063/5.0117329

McComas, W. F. (2024). Considering the lessons of curriculum studies in the design of science instruction: Varieties of meaning and implications for teaching and learning. *Education Sciences, 14*(3), Article 238. doi:10.3390/educsci14030238

Miao, Q. (2024). High-quality development of science education in rural schools: Value implications, existing challenges, and potential solutions. *Modern Distance Education*, *214*, 29–39.

Michael, W. B., Stanley, J. C., & Bolton, D. L. (1957). [Review of the book *Taxonomy of educational objectives, the classification of educational goals, handbook I: cognitive domain*, by B. S. Bloom]. *Educational and Psychological Measurement, 17*(4), 757–760. doi:10.1177/001316445701700420

Ministry of Education of the People's Republic of China. (2022). *Compulsory education science curriculum standards (2022 edition)*. Retrieved from http://www.moe.gov.cn

Mitchell, M. (2009). *Complexity: A guided tour*. Oxford, UK: Oxford University Press.

Nieminen, J. H., & Carless, D. (2023). Feedback literacy: A critical review of an emerging concept. *Higher Education, 85*(6), 1381–1400. doi:10.1007/s10734-022-00895-9

Pan, S. D. (2004). *From specialized to integrated science course: Research on the mode of science course offering in Chinese junior high school* (Doctoral dissertation). East China Normal University, Shanghai, China.

Pareek, R. B. (2019). An assessment of availability and utilization of laboratory facilities for teaching science at secondary level. *Science Education International, 30*(1).

Penuel, W. R., Reiser, B. J., McGill, T. A. W., Novak, M., Van Horne, K., & Orwig, A. (2022). Connecting student interests and questions with science learning goals through project-based storylines. *Disciplinary and Interdisciplinary Science Education Research, 4*(1), Article 1. doi:10.1186/s43031-021-00040-z

Piaget, J. (1997). *The principles of genetic epistemology*. London, UK: Routledge & Kegan Paul.

Scriven, M. (1991). Prose and cons about goal-free evaluation. *Evaluation Practice, 12*(1), 55–63. doi:10.1177/109821409101200108

Rakhmonkulov, F., & Usarov, S. (2019). Organization of practical and laboratory activities in the educational process. *European Journal of Research and Reflection in Educational Sciences, 7*(12).

Rios, F. A., & Stanton, C. R. (2011). *Understanding multicultural education: Equity for all students*. Lanham, MD: Rowman & Littlefield Education.

San Diego STEM Ecosystem. (2014). *A criteria for quality STEM/STEAM in San Diego*. Retrieved from https://www.sdstemecosystem.org/references/criteria-quality-stemsteam-san-diego

Schnitzler, K., Holzberger, D., & Seidel, T. (2021). All better than being disengaged: Student engagement patterns and their relations to academic self-concept and achievement. *European Journal of Psychology of Education, 36*(3), 627–652. doi:10.1007/s10212-020-00500-6

Shen, Y. (2024). Empowering digital intelligence: Creating new boundaryless learning spaces for rural primary schools—Research on the establishment and application of a digitized rural experimental science space. *Education and Equipment Research*, *2*, 68–98.

Soysal, Y. (2022). Science curriculum objectives' intellectual demands: A thematic analysis. *Journal of Science Learning, 5*(1), 127–140.

Spektor-Levy, O., Eylon, B.-S., & Scherz, Z. (2008). Teaching communication skills in science: Tracing teacher change. *Teaching and Teacher Education, 24*(2), 462–477. doi:10.1016/j.tate.2006.10.009

Stake, R. E. (1976). A theoretical statement of responsive evaluation. *Studies in Educational Evaluation*, *2*(1), 19–26.

State Council of the People's Republic of China. (2021). *Outline of the National Action Plan for Scientific Literacy (2021–2035)*. Retrieved from https://www.gov.cn/gongbao/content/2021/content_5623051.htm

Stufflebeam, D. L. (2003). The CIPP model for evaluation. In T. Kellaghan & D. L. Stufflebeam (Eds.), *International handbook of educational evaluation* (pp. 31–62). Dordrecht, Netherlands: Springer. doi:10.1007/978-94-010-0309-4_4

Suphasri, P., & Chinokul, S. (2021). Reflective practice in teacher education: Issues, challenges, and considerations. *PASAA: Journal of Language Teaching and Learning in Thailand, 62*, 236–264.

Taber, K. S. (2018). The use of Cronbach's alpha when developing and reporting research instruments in science education. *Research in Science Education, 48*, 1273–1296. doi:10.1007/s11165-016-9602-2

Thorndike, E. L. (1904). *Theory of mental and social measurements*. New York, NY: Teachers College, Columbia University.

Tyler, R. W. (1934). *Constructing achievement tests*. Columbus, OH: Ohio State University.

Vaidya, S. (2014). Curriculum evaluation. In S. Vaidya (Ed.), *Developing entrepreneurial life skills: Creating and strengthening entrepreneurial culture in Indian schools* (pp. 83–99). New Delhi, India: Springer. doi:10.1007/978-81-322-1789-3_6

Villa, G. (2018, October). *Bridging the gap between neuroscience and classroom teaching: New pathways in the application of social and emotional neuroscience* [Conference abstract]. 2nd International Conference on Educational Neuroscience, *Frontiers in Human Neuroscience*. doi:10.3389/conf.fnhum.2018.225.00013

Voss, P., Thomas, M. E., Cisneros-Franco, J. M., & De Villers-Sidani, É. (2017). Dynamic brains and the changing rules of neuroplasticity: Implications for learning and recovery. *Frontiers in Psychology, 8*, 1657. doi:10.3389/fpsyg.2017.01657

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.

Wang, R., Wang, Y., Xu, X., Li, Y., & Pan, X. (2023). Brain works principle followed by neural information processing: A review of novel brain theory. *Artificial Intelligence Review, 56*(Suppl 1), 285–350. doi:10.1007/s10462-022-10204-5

Wang, Y., Lavonen, J., & Tirri, K. (2019). Twenty-first century competencies in the Chinese science curriculum. In H. Liu, F. Dervin, & X. Du (Eds.), *Nordic-Chinese intersections within education* (pp. 151–171). Cham, Switzerland: Springer International Publishing. doi:10.1007/978-3-030-28588-3_7

Wang, Y. (2024). Practical exploration of integrated science curriculum development in junior high schools. *Primary and Secondary School Science Education*, *(2)*, 31–35.

Wanselin, H., Danielsson, K., & Wikman, S. (2022). Analysing multimodal texts in science—A social semiotic perspective. *Research in Science Education, 52*(3), 891–907. doi:10.1007/s11165-021-10027-5

Wieman, C. E. (2014). Large-scale comparison of science teaching methods sends clear message. *Proceedings of the National Academy of Sciences, 111*(23), 8319–8320.

Wu, X., Yang, Y., Zhou, X., Xia, Y., & Liao, H. (2024). A meta-analysis of interdisciplinary teaching abilities among elementary and secondary school STEM teachers. *International Journal of STEM Education, 11*(1), 38. doi:10.1186/s40594-024-00500-8

Yara, P., & Omondi, K. (2010). Teaching/learning resources and academic performance in mathematics in secondary schools in Bondo District of Kenya. *Asian Social Science, 6*(12), 126–132. doi:10.5539/ass.v6n12p126

Zhang, G., Zeller, N., Griffith, R., Metcalf, D., Williams, J., Shea, C., & Misulis, K. (2011). Using the context, input, process, and product evaluation model (CIPP) as a comprehensive framework to guide the planning, implementation, and assessment of service-learning programs. *Journal of Higher Education Outreach and Engagement*, *15*(4), 57–84.

Zhang, J. (2022). Strategies for experimental teaching and management of rural primary school science curricula. *Tianjin Education*, *9*, 162–163.

Zhang, R. (2019). *Research on the evaluation index system of STEM education projects in China* (Master's thesis). Beijing University of Posts and Telecommunications, Beijing, China.

Zhuo, Z. L. (2017). Beyond classroom: The implementation and evaluation of American K-12's out-of-school time STEM programs. *China Educational Technology*, *11*, 7–13.