Volume 1 Number 2 2025

iEducation

ISSN:3068-3769
https://ie.cscholar.com

iEducation

Volume 1, Issue 2, 2025

Quarterly (Issue No. 2)

Publishing Unit: ConnectSix Scholar Publishing INC

Publisher's website: http://www.cscholar.com/

Publisher's address:

6547 N Academy Blvd #2265

Colorado Springs CO 80918

US

Website of the *iEducation*:

https://ie.cscholar.com/

Table of contents

iEducation, 2025, 1(2), 1-14
https://doi.org/10.71204/hj1nf035

Fostering Social-Emotional Competencies Through English Instruction: A Case Study of SEL Integration in Chinese Primary Schools

Yijia Zhang [1], Zishu Meng [2,*]

[1] Queen Mary University of London, E1 4NS, United Kingdom

[2] Teacher's College Beijing Union University, Beijing 100011, China

*** Correspondence:**

Zishu Meng

sftzishu@buu.edu.cn

Received: 24 April 2025/ Accepted: 9 August 2025/ Published online: 13 August 2025

Abstract

Social-emotional learning (SEL) has gained global recognition for enhancing students' holistic development, yet its implementation in non-Western educational contexts remains underexplored. This mixed-methods case study investigates the integration of SEL principles into English language instruction at a Chinese primary school, addressing a significant gap in understanding how CASEL's five core competencies can support language acquisition in contexts where English is taught as a foreign language. Through in-depth interviews with three English teachers and systematic classroom observations, this study identified three primary challenges: student disengagement, proficiency disparities, and outdated instructional approaches. Following implementation of an integrated SEL-English curriculum based on CASEL's framework, classroom observations revealed significant improvements in student engagement, peer collaboration, and language participation. Specifically, incorporating self-awareness and self-management activities addressed attention difficulties, while social awareness and relationship-focused strategies mitigated proficiency gaps through structured peer support. This study contributes to both theoretical understanding of SEL application in subject-specific contexts and practical knowledge for language educators. The findings suggest that systematic integration of SEL principles can transform primary English instruction while supporting China's educational goals of cultivating well-rounded talents with both linguistic competence and social-emotional skills. The results have important implications for teacher education programs, curriculum development, and educational policy in emerging SEL contexts.

Keywords: Social-Emotional Learning (SEL); English Education; Primary School

1. Introduction

Social-Emotional Learning (SEL) has emerged as a cornerstone of educational reform globally, with particular significance for elementary education where foundational emotional competencies are established. Research indicates that early childhood, especially in lower primary grades, represents a critical period for emotional development that substantially influences future academic trajectories and overall well-being (Mah & Ford-Jones, 2012). Emotional education not only helps students develop positive affective attitudes but also enhances self-confidence and intrinsic motivation for learning. As China advances its educational modernization agenda in the 21st century, developing students' social-emotional competencies has become integral to national educational priorities that seek to align citizen development with broader societal goals (Chen, X & Hu, H, 2024). In this context, students must develop capabilities to navigate complex social and emotional challenges, with SEL offering a structured approach that aligns with both language acquisition goals and broader educational objectives.

This mixed-methods case study investigates two critical research questions: (1) What is the current state of English instruction at L Primary School with respect to social-emotional dimensions? and (2) How can the integration of social-emotional learning principles enhance English language teaching effectiveness in this context? To address these questions, in-depth interviews were conducted with three English teachers from grades one through three at L Primary School, systematically identifying key instructional challenges. Subsequently, an integrated SEL-English curriculum intervention was implemented and observed, with both student feedback and classroom dynamics analyzed through a structured observation protocol based on CASEL's five core competencies.

This study makes several significant contributions to the literature on SEL implementation and language instruction. First, it addresses the notable gap in research examining SEL integration in non-Western educational contexts, particularly in Chinese primary education where English is taught as a foreign language. Second, it offers empirically-based instructional strategies for language teachers seeking to incorporate social-emotional dimensions into their pedagogy. Third, it extends theoretical understanding of how CASEL's five core competencies can be operationalized within subject-specific teaching rather than as standalone curriculum components. These contributions have important implications for teacher education, curriculum development, and educational policy in contexts where SEL approaches are still emerging.

2. Literature Review

The integration of Social-Emotional Learning (SEL) into educational curricula has gained significant momentum globally, driven by substantial evidence supporting its positive impact on student development. This review synthesizes the existing literature by first establishing the robust empirical foundation for the benefits of SEL. It then narrows its focus to the specific integration of SEL within subject-specific pedagogy, particularly in language education. Subsequently, it examines the critical role of implementation fidelity and the cultural adaptation

required for SEL effectiveness in diverse contexts. Finally, by identifying key research gaps, this review establishes the rationale and contribution of the present study.

2.1. The Empirical Foundation and Broad Benefits of Social-Emotional Learning

A vast body of international research has firmly established the value of school-based SEL programs. A seminal meta-analysis synthesized findings from 213 programs involving over 270,000 students, provides the most compelling evidence to date. The study demonstrated that SEL interventions led to statistically significant improvements across multiple domains: students' social-emotional competencies (mean effect size = 0.57), attitudes toward self and others (mean effect size = 0.23), positive social behaviors (mean effect size = 0.24), and academic performance, with an average gain of 11 percentile points (mean effect size = 0.27).(J. Durlak et al., 2011) Further large-scale syntheses have reinforced these conclusions, establishing a global consensus that well-implemented SEL programs yield multidimensional benefits for students(John Payton et al., 2008).

2.2. The Integration of SEL into Language Education

Building on the general benefits of SEL, a growing area of inquiry explores its integration into specific academic subjects, including language education. The rationale is that embedding SEL into core instruction can create a synergistic effect, where emotional and social development enhances academic learning, and vice versa. In the context of language acquisition, competencies like self-awareness, self-management, and relationship skills are inherently linked to the processes of communication and cultural understanding. Scholars argue that enhanced emotional intelligence can lower the "affective filter," reducing language learning anxiety and thereby boosting participation and proficiency(Tevdovska, 2017).

For instance, researchers argued that integrating SEL into English instruction through specific strategies can bolster students' self-awareness and emotional regulation, which in turn improves language learning outcomes.(Bai et al., 2024) Similarly, within the Chinese context, Zhu (2022) developed a conceptual framework asserting that SEL integration can enhance three key dimensions of English learning: cultural awareness, self-development capacity, and social engagement. While these frameworks are compelling, empirical studies that test these proposed mechanisms and measure their direct impact on language proficiency remain an emerging area of research, signalling a need for more classroom-based investigations.

2.3. Implementation, Context, and Cultural Adaptation in China

While the potential of SEL is clear, its effectiveness is highly contingent on its adaptation to specific contexts. The non-Western context, particularly China, is not merely a different location but represents a distinct cultural and educational ecosystem that reshapes how SEL is understood and implemented. Unlike the individualistic underpinnings of many Western SEL models, Chinese education is deeply rooted in collectivist values and a strong tradition of moral education (deyu), which emphasizes collective harmony and social responsibility(An et al., 2021; Ping et al., 2004). Furthermore, the high-stakes pressure of the national college entrance examination (gaokao) often leads to a curriculum that heavily prioritizes academic knowledge over socio-emotional skills, posing a significant barrier to implementation.

Consequently, a direct transplantation of Western SEL frameworks is widely seen as ineffective. Instead, a process of cultural adaptation is underway. A prime example is the adaptation of the influential framework from the Collaborative for Academic, Social, and Emotional Learning (CASEL). Rather than simply adopting CASEL's five competencies, Chinese researchers and educators are actively mapping them onto the indigenous "Core Competencies" (hexin suzhi) framework promulgated by the Ministry of Education.

Despite this promising theoretical work, empirical research on the ground remains limited but growing. Qualitative studies, such as that by Chen (2024), have explored the implementation of school-based SEL curricula in China, revealing that while teachers recognize the value of SEL, they struggle to balance it with intense academic pressures. The few available quantitative studies, such as the one by Gong (2022) showing positive effects on social skills and engagement, are often small-scale (N=126) and conducted in well-resourced urban districts. This raises critical questions about the scalability and generalizability of such interventions across China's vast and heterogeneous educational landscape, highlighting a gap between conceptual adaptation and validated, large-scale implementation.

This synthesis reveals several critical gaps in the existing literature that the present study aims to address. Chief among them is the scarcity of empirical studies that implement and evaluate culturally-adapted models; while the adaptation of frameworks like CASEL to China's hexin suyang is discussed theoretically, its application in real-world classrooms remains under-researched. This is particularly true for subject-integrated strategies, as few studies have developed and assessed specific pedagogical practices for integrating SEL into primary school English instruction, despite scholars advocating for it as a feasible approach. Compounding these issues, the underlying process by which a culturally-adapted SEL model influences language engagement and proficiency within the unique pressures of the Chinese educational system is not yet well understood, highlighting a need for research that examines not only if these strategies work, but how they work in context.

3. Theoretical Framework: Conceptualizing SEL Integration in Primary English Education

This study operationalizes Social-Emotional Learning through the theoretical framework developed by the Collaborative for Academic, Social, and Emotional Learning (CASEL), established in 1994 (Du et al., 2019). CASEL's approach, significantly influenced by Goleman's emotional intelligence theory, conceptualizes SEL as a developmental process through which individuals acquire and effectively apply knowledge, attitudes, and skills necessary to understand and manage emotions, establish and achieve positive goals, feel and show empathy, maintain positive relationships, and make responsible decisions (Chang, 2020).

The research design is structured around CASEL's five core competencies, which provide both conceptual clarity and analytical utility for examining SEL integration in educational contexts. Each competency represents a distinct dimension of social-emotional development while simultaneously functioning as part of an interconnected system of skills and dispositions (Osher et al., 2016).

Self-awareness is the ability to recognize and understand one's own emotions, values, and strengths and weaknesses, and to maintain a realistic sense of self-efficacy and optimism. In language learning contexts, this includes students' awareness of their emotional responses to unfamiliar linguistic challenges and recognition of their language learning strengths and limitations.

Self-management refers to the ability to regulate one's emotions, thoughts, and behaviors in different situations, including stress management, impulse control, and goal setting. For language learners, this involves maintaining attention during instruction, persisting through difficulties in comprehension or production, and setting realistic language acquisition goals.

Social awareness is the capacity to understand and empathize with individuals from diverse backgrounds, recognize social and cultural norms, and identify available support networks and resources. In language education, this extends to cultural awareness and appreciation of diverse linguistic practices.

Relationship skills involve the ability to establish and maintain positive relationships, communicate effectively, listen actively, cooperate, resist peer pressure, negotiate conflicts, and seek help when needed. These skills are particularly relevant in communicative language teaching approaches that emphasize interactive learning.

Responsible decision-making is the ability to make informed choices by evaluating the potential consequences of actions, considering well-being, ethical standards, safety, and social norms. For language learners, this includes making appropriate choices about language use in varied social contexts.

This framework provides analytical categories for examining the challenges faced in primary English education and structuring the development and implementation of integrated SEL-English instructional approaches. By mapping observed classroom dynamics and teaching strategies to these five competencies, the study can systematically analyze both existing practices and potential areas for enhancement.

4. Research Design

4.1. Research Questions and Objectives

This study investigates two primary research questions:

What social-emotional dimensions are evident in current English instruction at L Primary School, and what challenges do teachers face in addressing students' social-emotional needs?

How can systematic integration of SEL principles enhance English language teaching effectiveness and student engagement in this context?

These questions address the identified research gap regarding SEL implementation in Chinese language education contexts and aim to generate both theoretical insights and practical strategies for teacher education.

4.2 Research Methods

This study employed a qualitative case study methodology, incorporating multiple procedures to ensure research rigor and trustworthiness. The research was conducted at a public primary school in Beijing. The school is located in an urban area and serves a student population with a relatively diverse but generally middle-class socio-economic background. Each class typically consists of around 40 students, which is a common class size in urban Chinese primary schools. The participating teachers had varying levels of teaching experience, ranging from 4 to 20 years, and all held professional teaching certifications. Before the research, teachers are provided integrated SEL-English curriculum based on CASEL's framework and familiar with the curriculum. Data collected from semi-structured interviews and systematic classroom observations were triangulated to enhance the validity of the findings. Furthermore, potential researcher bias was mitigated through peer debriefing, and inter-coder reliability was established to ensure consistency in the data analysis process. This methodological framework is consistent with prior investigations of Social and Emotional Learning (SEL) implementation and is specifically adapted to the context of primary education in China.

Data collection proceeded in two sequential phases. In Phase One, in-depth semi-structured interviews were conducted with three English teachers (grades 1-3) at L Primary School (see Table 1 for participant demographics). The interview protocol comprised 11 questions systematically aligned with CASEL's five core competencies: self-awareness (questions 1-2), self-management (questions 3-4), social awareness (questions 5-6), relationship skills (questions 7-8), and responsible decision-making (question 9). Two additional questions addressed general teaching challenges and resource needs. Interviews ranged from 60-90 minutes, were audio-recorded with participant consent, and transcribed verbatim for analysis. Member checking was employed to ensure transcription accuracy.

In Phase Two, an integrated SEL-English curriculum informed by challenges identified in Phase One was implemented and observed. Classroom observations employed a structured protocol adapted from two validated instruments: Yoder's (2014) Teacher Social-Emotional Teaching and Ability Self-Assessment Tool and Gong's (2022) SEL observation scale. Following content validation by three educational experts specializing in SEL, the original ten-indicator framework was refined to nine indicators that specifically addressed the research questions. The "balanced instruction" indicator was eliminated due to limited relevance to SEL dimensions based on expert feedback. Inter-rater reliability was established through parallel observations with two trained observers (Cohen's kappa = 0.87).

Table 1. Interview Information

Interviewee	Subject	Title	Teaching Age	Interview Time
Teacher1	English	Intermediate teacher	20 years	1.5 hour
Teacher2	English	Intermediate teacher	8 years	1hour
Teacher3	English	Intermediate teacher	4 years	1hour

The refined observation framework includes indicators designed to create a positive and supportive learning environment, such as "Teacher language," "Warmth and support," and "Academic press and expectations." Additionally, "Responsibility and choice" aligns with SEL's focus on responsible decision-making, while "Cooperative learning" and "Classroom discussions" correspond to SEL's emphasis on relationship skills. Indicators like "Self-assessment and self-reflection" and "Student-centered discipline" are linked to the self-management domain of SEL, while "Competence building" addresses both academic and SEL skill development.

5. Results: Examining SEL Dimensions in Primary English Instruction

5.1. Thematic Analysis of Teacher Interviews

Analysis of interview data revealed three consistent themes across all participants: (1) recognition of limitations in traditional instructional approaches, (2) identification of student behavioral and engagement challenges, and (3) intuitive but unsystematic application of SEL-aligned strategies to address these challenges. The following sections examine these themes through the lens of CASEL's five core competencies, illustrated with representative excerpts from participant interviews.

5.1.1. Self-Awareness and Self-Management Challenges

Self-management challenges emerged as a primary concern across all grade levels. Teacher 1 (Grade 1) reported significant engagement difficulties: "In each of my six classes, I consistently observe 4-5 students who cannot maintain attention for more than 10 minutes. They sleep, draw, or play with items in their desks." This observation illustrates the self-regulation difficulties that young language learners experience, particularly in a context where English represents an unfamiliar linguistic and cultural system. A specific case highlighted by Teacher 1 involved a student with physical disabilities who demonstrated significant impulse control challenges: "He frequently leaves his seat during instruction, approaches the teacher's desk, and interrupts ongoing activities." These behaviors reflect difficulties with the self-management competencies of impulse control and emotional regulation, which CASEL identifies as foundational for academic engagement.

Teacher 3 highlighted the increasing prevalence of psychological challenges that affect self-awareness: "I've noticed more students struggling with anxiety about English learning. One particular student in my class constantly hesitates to answer questions because she's afraid of making mistakes. When she does make an error, she becomes extremely self-critical and sometimes refuses to participate further." This example illustrates deficits in emotional self-awareness and accurate self-perception, which impedes language learning progress. In response, Teacher 3 reported adapting her approach: "I've learned to use a gentler tone with this student and provide more affirmation even for partial answers. This seems to reduce her anxiety somewhat, but I wish I had more structured strategies to help build her confidence."

5.1.2. Social Awareness and Relationship Skills

Interview data suggested that relationship skills and social awareness presented both challenges and opportunities in the English classroom. Teacher 2, who taught third-grade students, noted a considerable disparity in English proficiency within her class: "Some students use expressions well beyond our curriculum, while others struggle with basic vocabulary we covered months ago. This creates tension during pair work or group activities because stronger students become impatient with peers who can't keep up." This observation highlights how proficiency disparities can complicate the development of relationship skills such as cooperation and communication.

However, teachers also recognized the potential of leveraging relationship skills to enhance learning. Teacher 2 reflected: "When I organize activities where stronger students help teach concepts to others, I see improvements in both groups. The helpers develop patience and communication skills, while the struggling students often understand better from peers than from me." This insight suggests an intuitive understanding of how social dynamics can support language learning, though teachers reported lacking systematic approaches to structuring these interactions for optimal outcomes.

5.1.3. Responsible Decision-Making

The interviews revealed limited explicit attention to responsible decision-making in current English instruction. When asked about how they help students make responsible choices, teachers primarily focused on behavioral compliance rather than ethical reasoning or consequential thinking. Teacher 1 noted: "I mainly emphasize following class rules and completing assignments on time. We don't have much opportunity to address deeper decision-making in our curriculum." This suggests a potential area for growth in SEL integration, as responsible decision-making extends beyond compliance to include critical thinking about language choices and cultural implications.

5.2. Challenges in English Teaching

The data revealed three principal challenges in primary English instruction at L Primary School, each with distinct social-emotional dimensions:

5.2.1. Self-Awareness and Self-Management Challenges

The data revealed significant challenges related to students' self-awareness and self-management capabilities that directly impacted English learning. Teachers reported that approximately 15-20% of students across all grade levels demonstrated consistent difficulties in maintaining attention during English instruction. Teacher 1 noted that "in each of my six classes, one or two students regularly fall asleep during lessons," indicating challenges with self-awareness of learning engagement. Teacher 2 provided quantitative evidence of self-management difficulties, reporting that "while homework completion rates in most subjects range from 85% to 90%, English assignment completion in some classes drops to approximately 70%." This disparity suggests that students struggle more with self-management specifically in the English language learning context.

The interviews uncovered several factors contributing to these self-management challenges. First, as English is not students' first language, they must build language skills from fundamental levels, requiring substantial cognitive resources and sustained attention. Second, some students develop simplified approaches to language learning, viewing it primarily as vocabulary memorization. When the curriculum advances to more complex communicative tasks, these students experience frustration and disengagement. Teacher 3 observed: "Many students who performed well in earlier grades begin to struggle when we move from vocabulary drills to actual communication. They don't have strategies to manage this transition, and their frustration often manifests as disruptive behavior or withdrawal."

5.2.2. Proficiency Disparities and Social Dynamics

All three teachers identified significant proficiency disparities within their classrooms, which created challenges for both instruction and peer relationships. Teacher 2 reported that in the third-grade final English exam, while 97% of students passed, 21% received C grades, and only 16% achieved full marks. These disparities created social dynamics that often reinforced existing gaps: high-performing students tended to interact primarily with each other during English activities, while struggling students became increasingly marginalized.

The teachers identified two primary factors contributing to these disparities. First, home environment played a crucial role, with students from more supportive and resource-rich home environments demonstrating stronger English performance. Second, cognitive differences among students led to varied aptitudes for language learning. As Teacher 1 explained: "Some students naturally excel in language acquisition while struggling with subjects like mathematics. Others show the opposite pattern. These cognitive differences become apparent in the classroom and can significantly impact students' engagement with English learning."

5.2.3. Instructional Approaches and Teacher Development

The interview data revealed that instructional approaches constituted a third significant challenge. All three teachers acknowledged limitations in current teaching methods, with Teacher 3 noting that "some colleagues continue to rely almost exclusively on textbook recitation and traditional grammar exercises, with minimal interactive components." This observation highlights a gap between contemporary language teaching theories, which emphasize communicative competence, and actual classroom practices that often remain teacher-centered and text-focused.

Teachers also reported limited professional development opportunities specifically focused on integrating social-emotional dimensions into language instruction. Teacher 2 reflected: "We receive general pedagogical training, but very little guidance on addressing the emotional aspects of language learning or creating emotionally supportive classroom environments. When students experience anxiety or frustration with English, I mostly rely on intuition rather than specific strategies." This suggests a need for targeted professional development that bridges language pedagogy with social-emotional learning principles.

6. Discussion

6.1. Enhancing Student Interest Through Self-Awareness and Self-Management

The findings reveal significant potential for addressing student engagement challenges through systematic integration of self-awareness and self-management dimensions of SEL. The observed disengagement patterns-including students sleeping during class, completing other subjects' work during English instruction, and displaying anxiety about participation-directly correspond to deficits in self-awareness (recognizing one's emotional state during learning) and self-management (regulating attention and behavior). This aligns with Gong's (2022) findings that SEL-integrated activities significantly enhance student engagement through increased self-regulatory capacity, though the current study extends this work by specifically examining the language learning context.

The classroom implementation phase of the study demonstrated that incorporating creative and interactive activities specifically designed to develop self-awareness and self-management skills had measurable effects on student engagement. When reflective journaling about emotions experienced during English learning (self-awareness) and specific attention-focusing techniques (self-management) were introduced, increased on-task behavior and participation rates were observed. As Teacher 1 noted after implementing these strategies: "The reflective activities helped students recognize when they were becoming disengaged or frustrated, which seemed to help them regulate these responses more effectively."

Additionally, expanding students' cultural awareness through culturally responsive content enhanced both motivation and self-awareness. For example, when addressing traditional festivals, teachers incorporated proverbs or expressions related to these festivals in both English and Chinese. This approach helped students connect new language content to familiar cultural contexts, making the learning process more accessible and personally meaningful. This finding aligns with previous research by Zhu (2022), who found that culturally integrated SEL approaches enhanced students' engagement with foreign language content.

6.2. Addressing Proficiency Disparities Through Social Awareness and Relationship Skills

The pronounced proficiency disparities observed in this study (with 21% of Grade 3 students receiving C grades despite a 97% pass rate) reflect challenges that extend beyond conventional pedagogical approaches. The findings suggest that social awareness and relationship skills-two key SEL competencies-offer promising avenues for addressing these disparities. Through collaborative learning structures where more proficient students support peers, teachers can leverage social dynamics to create what Vygotsky (1978) termed "zones of proximal development." This approach was exemplified when Teacher 2 implemented mixed-ability grouping with stronger students as "language ambassadors," resulting in increased participation from previously disengaged students.

During the implementation phase, it was observed that strategic grouping based on varied proficiency levels, combined with explicit instruction in peer coaching techniques, significantly enhanced learning outcomes for struggling students. For example, when stronger students were taught specific ways to provide feedback and encouragement to peers (relationship skills) and to

recognize signs of confusion or frustration in others (social awareness), collaborative activities became more productive for all participants. As one student remarked: "I understand better when my classmate explains it, and she doesn't make me feel bad when I make mistakes."

These findings extend previous research by Bai et al. (2021), who found that SEL-integrated language instruction enhanced peer relationships. However, this study specifically demonstrates how structured approaches to developing relationship skills and social awareness can directly address the challenge of proficiency disparities in Chinese primary English education contexts.

6.3. Improving Teacher Professional Development

The findings highlight the critical role of teacher self-awareness and social awareness in effective SEL integration. The observation that all three participating teachers intuitively incorporated some SEL elements without systematic understanding suggests that teacher education programs should explicitly develop teachers' metacognitive awareness of these practices. As Schon's (1983) concept of "reflection-in-action" suggests, teachers who develop heightened awareness of their own social-emotional practices can more effectively model and teach these competencies. The practice at L Primary School of requiring at least 20 peer observations per semester represents a promising structural support for developing this reflective capacity, though the data suggest these observations would benefit from explicit SEL-focused protocols.

The implementation phase revealed that teachers benefit from specific professional development in three areas: (1) recognizing and responding to students' emotional states during language learning, (2) designing activities that intentionally develop specific SEL competencies while addressing language objectives, and (3) managing their own emotional responses to challenging classroom situations. This third area proved particularly important, as Teacher 1 noted: "I realized that my own frustration with disengaged students was often visible to the class and actually reinforced negative patterns. Learning to regulate my emotional expressions has created a more positive atmosphere."

These findings align with research by Li et al. (2023), who emphasized the importance of teacher emotional competence in successful SEL implementation. However, the current study extends this work by identifying specific professional development needs for language teachers in the Chinese context, where traditional teacher-centered pedagogies remain influential despite curriculum reforms emphasizing student-centered approaches.

6.4. Future Research Directions

This study reveals several promising avenues for future research on SEL integration in language education contexts. First, longitudinal studies examining the sustained impact of SEL-integrated language instruction on both language proficiency and social-emotional development would provide valuable insights into long-term outcomes. Second, comparative studies across different grade levels, socioeconomic contexts, and geographical regions within China would help identify contextual factors that influence SEL implementation effectiveness. Third, research specifically examining how different aspects of language acquisition (e.g., vocabulary

development, oral communication, reading comprehension) interact with specific SEL competencies could inform more targeted integration approaches.

Additionally, future research should investigate how teacher education programs and professional development initiatives can most effectively prepare language teachers to integrate SEL principles into their instruction. This includes examining both pre-service teacher education curricula and in-service professional development models to identify optimal approaches for developing teachers' capacity to implement SEL-integrated language instruction.

7. Conclusions

This study examined the integration of Social-Emotional Learning principles into English language instruction at a Chinese primary school, addressing the notable gap in research on SEL implementation in non-Western educational contexts. The findings demonstrate that significant challenges in primary English instruction - including student disengagement, proficiency disparities, and outdated teaching methodologies - can be effectively addressed through systematic integration of CASEL's five core competencies. Specifically, the development of self-awareness and self-management skills appears to enhance student engagement with language learning, while social awareness and relationship skills offer promising approaches for addressing proficiency disparities through collaborative learning structures.

The theoretical and practical contributions of this study are threefold. First, it extends SEL research into the specific context of Chinese primary English education, demonstrating both the cultural adaptability and subject-specific applications of CASEL's framework. Second, it provides empirically-grounded strategies for teachers seeking to integrate social-emotional dimensions into language instruction, particularly in contexts where English represents a foreign language. Third, it highlights the importance of teacher self-awareness and reflective practice in successful SEL implementation, suggesting implications for teacher education programs and professional development structures.

Several limitations warrant consideration when interpreting these findings. First, the case study methodology employed at a single primary school limits generalizability, particularly given China's diverse educational contexts. Second, the relatively short implementation period restricts the ability to assess long-term impacts of SEL integration on language proficiency and social-emotional development. Third, the absence of quantitative measures of student outcomes precludes statistical analysis of intervention effectiveness. Future research should address these limitations through multi-site comparative studies, longitudinal designs that track developmental trajectories, and mixed-methods approaches that incorporate standardized measures of both language proficiency and social-emotional competencies(Zhu, 2022).

Despite these limitations, this study makes a significant contribution to understanding how SEL principles can enhance primary English education in China. As the country continues to prioritize both English language proficiency and holistic student development, SEL integration offers a promising approach that addresses both objectives simultaneously. By developing students' social-emotional competencies within the context of language instruction, educators can foster not

only linguistic skills but also the emotional intelligence and interpersonal capabilities that are increasingly valued in contemporary society.

Author Contributions:

Conceptualization, Yijia Zhang and Zishu Meng; methodology, Yijia Zhang; software, Yijia Zhang; validation, Yijia Zhang and Zishu Meng; formal analysis, Yijia Zhang; investigation, Yijia Zhang; resources, Yijia Zhang; data curation, Yijia Zhang; writing-original draft preparation, Yijia Zhang; writing-review and editing, Yijia Zhang and Zishu Meng; visualization, Yijia Zhang; supervision, Zishu Meng; project administration, Zishu Meng; funding acquisition, Zishu Meng. All authors have read and agreed to the published version of the manuscript.

Funding:

This research received no external funding.

Institutional Review Board Statement:

Ethical review and approval were waived for this study due to its nature as an educational, non-interventional study conducted in accordance with local regulations in China. No personal or sensitive data were collected, and all participants provided informed consent voluntarily.

Informed Consent Statement:

Informed consent was obtained from all subjects involved in the study. No personally identifiable information has been disclosed.

Data Availability Statement:

The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author(s).

Acknowledgments:

The authors would like to express sincere gratitude to the English teachers at L Primary School for their cooperation and participation in the interviews and classroom observations. This study would not have been possible without their generous contributions. Special thanks to all colleagues and peers who provided insightful feedback during the preparation of this manuscript.

Conflict of Interest:

The authors declare no conflict of interest.

References

An, L., Vaid, E., Elias, M. J., Li, Q., Wang, M., & Zhao, G. (2021). Promotion of social and emotional learning in a Chinese elementary school. Social Behavior and Personality: An International Journal, 49(10), 1–9.

Bai, B., Shen, B., & Wang, J. (2024). Impacts of social and emotional learning (SEL) on English learning achievements in Hong Kong secondary schools. Language Teaching Research, 28(3), 1176–1200.

Chang, G. X. (2020). Whether Social Emotional Learning (SEL) independent courses are effective. QingDao University.

Chen, L. (2024). Strategies for Implementing Social and Emotional Learning (SEL) in education. Applied & Educational Psychology, 5(4), 48–53.

Chen, X & Hu, H. (2024). Chinese-style modernization and comprehensive human development: An analysis from the perspective of the sinicization and modernization of Marxism. 4(13), 1–11.

Du, Y., & Mao, Y. Q. (2019). From specialized courses to comprehensive reform: Strategy model transformation for promoting Social Emotional Competencies. Global Education, 48(05), 39–53.

Gong, Y. Y. (2022). An exploration of the integration of Social and Emotional Learning with subject teaching. Shanghai Normal University.

J. Durlak, R. Weissberg, Allison B. Dymnicki, Rebecca D. Taylor, & Kriston B. Schellinger. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. Child Development.

John Payton, R. Weissberg, & Molly Pachan. (2008). The positive impact of Social and Emotional Learning for kindergarten to eighth-grade students: Findings from three scientific reviews. Technical Report.

Li et al. (2023). Social and Emotional Competence in adolescents: Concepts, development, and practice. Journal of Gannan Normal University, 44(05), 104–111.

Mah, V. K., & Ford-Jones, E. L. (2012). Spotlight on middle childhood: Rejuvenating the 'forgotten years'. Paediatrics & Child Health, 17, 81–83.

Osher, D., Kidron, Y., Brackett, M., Dymnicki, A., Jones, S., & Weissberg, R. P. (2016). Advancing the science and practice of Social and Emotional Learning: Looking back and moving forward. Review of Research in Education, 40(1), 644–681.

Ping, L., Minghua, Z., Bin, L., & Hongjuan, Z. (2004). Deyu as moral education in modern China: Ideological functions and transformations. Journal of Moral Education, 33(4), 449–464.

Schon, D.A. (1983). The reflective practitioner: How professionals think in action. Basic Books.

Tevdovska, E. S. (2017). The impact of Emotional Intelligence in the context of language learning and teaching. SEEU Review, 12(1), 125–134.

Yoder, N. (2014). Self-assessing Social and Emotional Instruction and Competencies: A tool for teachers. Center on Great Teachers and Leaders. https://eric.ed.gov/?id=ED553369

Zhu, T. (2022). Elementary school English unit integral teaching based on social emotional learning. GuangDongJiaoYu, 8, 31–32.

iEducation, 2025, 1(2), 15-29
https://doi.org/10.71204/cq3kn353

Assessment of Reading Literacy in PISA-D: Connotation, Framework Design and Enlightenment

Xinyi Tian [1,*]

[1] Beijing Language and Culture University, Beijing 100083, China,

*** Correspondence:**

Xinyi Tian

txyashore@163.com

Received: 16 June 2025/Accepted: 15 August 2025 /Published online: 24 August 2025

Abstract

To assist low- and middle-income countries (LMICs) in participating in international education assessment, diagnosing their education systems, and advancing education-related Sustainable Development Goals, the PISA-D project has achieved three key innovations based on PISA. Firstly, it expanded assessment targets to include both in-school students and out-of-school youth through categorized design, filling the gap in traditional assessments that neglect out-of-school populations. Secondly, it developed a "downward-compatible" reading literacy assessment framework, adding the lower-difficulty Level 1c and optimizing item difficulty distribution to accurately reflect students' performance in LMICs. Thirdly, it designed contextual questionnaires based on Willms' prosperity model, refining equity and equality indicators to comprehensively capture deep-seated factors influencing academic performance. These innovations provide significant references for China to improve compulsory education Chinese subject assessment, particularly in enhancing the sensitivity of assessing vulnerable groups and optimizing the regional education quality monitoring system.

Keywords: PISA-D; Reading Literacy; Prosperity Model; Assessment of Compulsory Education Quality

1. Introduction

In 2014, drawing on the experience of 40 middle-income countries and 4 low-income countries participating in the Programme for International Student Assessment (PISA) since 2000, the Organization for Economic Co-operation and Development (OECD) launched the PISA for Development (PISA-D) initiative, a development assistance program. The initiative aims to enable broader participation of low- and middle-income countries (LMICs) in international education assessment through three strategic approaches: expanding the assessment population, enhancing the reading literacy assessment tools, and refining the relevant factor questionnaires. By doing so, PISA-D seeks to assist these countries in comparing educational opportunities and

outcomes within their own systems and those of other countries, strengthening their capacity for evidence-based education policy-making and feedback mechanisms, striving to reduce regional disparities, mitigating the influence of socioeconomic status, and ultimately improving learning outcomes.

The United Nations Educational, Scientific and Cultural Organization (UNESCO) proposed a new global education goal in its 2030 Agenda for Sustainable Development—Sustainable Development Goal 4 (SDG 4)—which aims to ensure that by 2030, all girls and boys complete free, equitable, and quality primary and secondary education, acquiring relevant and effective learning outcomes; and ensure that all youth and a substantial proportion of adults, both men and women, achieve literacy and numeracy (UNESCO, 2016). The Agenda also designates PISA assessment results as a key metric for global basic education quality. Consequently, PISA is expected to include more UN member states in the future, extending its influence to middle-income countries outside the OECD. The OECD report Experience of Middle-Income Countries with PISA 2000–2015 reveals that, except for Vietnam, 15-year-old students in middle-income countries participating in PISA have academic performance lower than their OECD counterparts, with significant variability in performance among these middle-income countries. In many low- and middle-income countries (LMICs), high out-of-school rates among school-aged children mean that many 15-year-olds are ineligible for PISA testing, resulting in a national population coverage rate of less than 50% for tested 15-year-olds in LMICs (Lockheed et al., 2015). Regarding the contextual factors influencing academic performance, PISA's measurement of family socioeconomic status (SES) indicators does not fully capture the true impact of negative factors such as parental education levels, income, and poverty on academic underperformance in LMICs. These issues have compromised the scientific validity and applicability of test results in LMICs to some extent, reducing these countries' motivation to participate in PISA.

As the number of countries participating in PISA continues to grow, PISA's assessment framework and implementation model require further optimization to meet the more diverse needs of participating nations. In response, the OECD launched the six-year PISA for Development (PISA-D) project, which involves the categorized design of assessment tools and contextual factor questionnaires for in-school students and out-of-school youth. Specifically, the in-school assessment targets 15-year-old students enrolled in Grade 7 or above, consistent with PISA's original scope. The out-of-school youth assessment covers 14–16-year-olds not included in the in-school assessment, including both dropouts and students enrolled in Grade 6 or below.

PISA-D involves countries and regions across Asia, Africa, and Latin America, including Cambodia, Ecuador, Guatemala, Honduras, Lao PDR, Paraguay, Senegal, and Zimbabwe. These participants collaborated in developing assessment tools and contextual questionnaires tailored to their educational contexts.

The initiative aims to enable broader participation of LMICs in international education assessment through three strategic approaches: expanding the assessment population, enhancing the reading literacy assessment tools, and refining the relevant factor questionnaires. By doing so, PISA-D seeks to assist these countries in comparing educational opportunities and outcomes within their own systems and those of other countries, strengthening their capacity for evidence-

based education policy-making and feedback mechanisms, striving to reduce regional disparities, mitigating the influence of socioeconomic status, and ultimately improving learning outcomes.

2. The Implementation Purpose of PISA-D

2.1. The Impact of Out-of-School Youth on Pisa Test Results

Notably, the proportion of 15-year-old out-of-school students exerts a significant impact on PISA test results, as evidenced by substantial research. In many developing countries, a high proportion of 15-year-olds are out of school due to multiple economic and social factors (UNESCO, 2025). Take Vietnam as an example: in the PISA 2015 assessment, the Vietnamese sample covered only 48.5% of the country's 15-year-olds, the lowest coverage rate among all participating countries(OECD, 2016a) . OECD analyzed that the 15-year-olds not included in the sample were likely academically weaker than those tested. By comparing Vietnam's science scores with other PISA data (assuming the scores of out-of-school 15-year-olds were at or below the national median level), he estimated that Vietnam's "real" scores might be 50 to 60 points lower than its average science scores, which would mean Vietnam's ranking in the science subject could drop from 8th to 35th - 40th (Lockheed et al., 2015).

Similar situations are also observed in other countries. In some low- and middle-income countries, a large number of 15-year-olds cannot access education due to poverty and a shortage of educational resources (Lockheed et al., 2015). When these countries participate in PISA assessments, if only in-school 15-year-olds are taken as samples, the test results may overestimate the overall level of the country's students. This is because out-of-school students, often lacking systematic education, are obviously backward in knowledge and skills compared with in-school students. If this group of students is taken into account, the overall scores will inevitably be affected, pulling down the country's average scores and related rankings in PISA assessments.

2.2. The Selection Method of PISA-D Out-of-School Youth

In the PISA-D project, the selection of out-of-school youth is a complex and rigorous process. Firstly, the scope of potential out-of-school youth groups is determined through multi-channel data collection. On the one hand, statistical data from educational authorities of various countries are used to understand the dropout situation of students in different regions and schools, and to obtain preliminary lists or relevant information of out-of-school students. For instance, some countries regularly calculate the dropout rates of students in compulsory education and record information such as the time and reasons for students leaving school, which provides an important basis for determining the scope of out-of-school youth (OECD, 2020a).

On the other hand, cooperation is carried out with community organizations and non-governmental organizations to utilize the information they have acquired in their grassroots work. These organizations come into contact with a large number of out-of-school youth who have dropped out due to poverty, family changes and other reasons while carrying out poverty alleviation, educational assistance and other work, and they can provide more detailed and

accurate information about out-of-school youth, including students' family conditions and where they go after dropping out of school.

After identifying the potential groups, a combination of stratified sampling and random sampling is adopted to select samples. Stratification is conducted according to factors such as the economic development level and educational resource status of different regions. In areas with backward economies, scarce educational resources and high dropout rates, the sampling proportion is appropriately increased to ensure that out-of-school youth in these areas have sufficient representativeness. Then, random sampling is carried out within each stratum to select a certain number of out-of-school youth as the final samples. For example, in a certain poverty-stricken area, after preliminary data statistics, there are 500 out-of-school youth, and according to the sampling plan, 50 are determined to be selected for research. 50 are selected from these 500 through methods such as random number generation to ensure that each out-of-school youth has the same probability of being selected.

Meanwhile, to ensure the comprehensiveness and accuracy of the samples, cross-validation and supplementation of the samples are conducted. Data obtained from educational authorities and community organizations are compared to check for omissions or duplicates. For some remote areas or special groups where accurate information is difficult to obtain, supplementary sampling is carried out through on-site visits, volunteer surveys and other methods, so as to make the selected samples of out-of-school youth truly reflect out-of-school groups with different backgrounds and situations as much as possible, and provide reliable data support for the subsequent research and evaluation of out-of-school youth in the PISA-D.

Based on the above analysis, PISA-D aims to establish a more inclusive and equitable international student assessment framework to support evidence-based policy-making in low- and middle-income countries (LMICs) and assist them in assessment progress toward achieving the 2030 Education Sustainable Development Goals (SDGs). The project’s evaluation objectives are specifically reflected in the following three aspects:

First, by expanding the assessment scope to include both in-school students and out-of-school youth in participating LMICs, enhancing the sensitivity of literacy assessment tools, and developing targeted contextual factor questionnaires and data collection instruments, PISA-D seeks to comprehensively capture disparities in student learning, teacher instruction, and school education system operations in LMICs. This provides a platform for these countries to diagnose their education systems, compare the strengths and weaknesses of educational systems, and identify areas for improvement. Meanwhile, maintaining the comparability of assessment tools and questionnaires with PISA enables in-depth analysis of how contextual factors across different backgrounds correlate with student academic achievement, offering policy recommendations to education decision-makers in participating countries on how to support better student learning, more proficient teaching, and more effective school system operations.

Second, PISA-D designs learning and capacity-building programs to encourage and assist LMICs interested in and motivated to participate in PISA (OECD, 2018a). Through a series of safeguards, the project helps these countries build the capacity to conduct large-scale academic

quality assessment, analyze, and utilize assessment results to inform national policies and evidence-based decision-making.

Finally, the implementation of PISA-D contributes to the realization of education SDGs. The project's development experience and assessment items will be integrated into PISA 2022, while the out-of-school youth assessment will become an optional component of future PISA cycles. PISA-D aims to enable more LMICs to benefit from scientific and effective international comparative data for validating national education policies, improving educational quality, and assessment progress toward education SDGs (OECD, 2018a).

3. The "Downward-Compatible" Reading Literacy Assessment of PISA-D

3.1. The Assessment Framework and Proficiency Level of Reading Literacy

PISA-D defines reading literacy as an individual's ability to understand, use, reflect on, and engage with written texts to achieve personal goals, develop knowledge and potential, and participate in society, which is largely consistent with PISA 2009's definition of reading literacy. The development of PISA-D reading literacy is based on three main task characteristics:

Texts: Refers to the scope of reading materials, involving text media, text structure, text form, and text type.

Situation: Refers to the context or purpose of reading, covering personal, educational, vocational, and public scenarios.

Processes: Refers to the cognitive methods, strategies, and purposes through which test-takers engage with texts, including accessing and retrieving information, integrating and interpreting content, and reflecting and evaluating.

Given that the assessment population includes not only students across multiple grades but also out-of-school youth in the labor market, PISA-D test items address a broader range of reading contexts that students may encounter both in and out of school—such as job hunting through classified ads in newspapers or following workplace instructions. In the dimension of reading processes, PISA-D emphasizes basic reading skills by introducing a "literal comprehension" process, requiring students to understand the explicit literal meanings of individual words, sentences, and paragraphs. It also expands the "information retrieval" process, expecting students to locate explicitly stated single pieces of information (e.g., individual words or phrases) (UNESCO, 2018).

Within the original PISA reading literacy assessment framework, PISA-D appropriately lowers the assessment standards and tilts test items toward lower proficiency levels to increase the detection rate of students in LMICs who perform at the bottom in reading. Meanwhile, PISA-D retains sufficient items for equating with PISA to ensure the comparability of results. Additionally, given the historically low reading performance of students in LMICs on PISA, approximately 65% of PISA-D reading literacy test items are at Level 2 or below in difficulty distribution to better reflect these countries' students' academic performance levels (OECD, 2018a). In terms of

dimensional distribution, the processes of integration and interpretation in reading and educational situation in reading account for the largest proportions, as shown in table 1.

Table 1. Reading item counts by framework category

Process	**Items**	**Percent**
Access and retrieve	22	33%
Integrate and interpret	31	46%
Reflect and evaluate	14	21%
Situation	**Items**	**Percent**
Personal	22	33%
Educational	21	31%
Occupational	4	6%
Public	2	30%

* OECD (2018), PISA for Development Assessment and Analytical Framework: Reading, Mathematics and Science, OECD Publishing, Paris, 10.

PISA-D identifies task or text characteristics closely related to reading difficulty. The first is quantitative characteristics and conditions, referring to how many elements readers must locate or consider in the text to answer a question. Tasks are simpler when fewer characteristics or conditions are required. The second is proximity of required information segments, i.e., how close related information segments are to each other in the text. The third is amount of competing information, defined as how much information in the text might be mistakenly identified by readers as target information; more competing information makes tasks more difficult. The fourth is salience of necessary text information, indicating how easily readers can locate required information. Target information is more easily found when it is clearly signaled by headings, positioned near the text's beginning, or part of very short texts. The fifth is relationship between task and required information, referring to the complexity of the relationship between the task and the text. Tasks are more difficult if their wording is linguistically complex or requires readers to infer their relationship to the text through reasoning. The sixth is semantic match between task and text, measuring the degree to which the task uses the same words as relevant parts of the text or words from the same lexical domain; closer lexical alignment simplifies tasks. The seventh is familiarity of information required to answer, i.e., readers' familiarity with the task's content or theme. The eighth is discourse structure, encompassing the form and syntactic complexity of the discourse. The ninth is degree of extra-textual information required, referring to the extent to which readers must use prior knowledge to answer questions. Tasks are more challenging when they require readers to reflect on and evaluate the text or construct meaning using prior knowledge to supplement the text.

PISA 2015 classifies reading literacy assessments into seven proficiency levels (1b–6), with Level 6 being the highest and Level 2 marking the baseline level of student reading performance, serving as one of the metrics for youth literacy under SDG 4. Students reaching this level are considered to have begun developing the ability to effectively participate in social life as students, workers, and citizens (OECD, 2016b). Based on an analysis of task or text characteristics, PISA-D implements a "downward-compatible" adaptation of PISA's reading performance proficiency levels, placing in-school and out-of-school youth assessment participants on a unified proficiency scale to present a comprehensive picture of academic performance among LMIC students.

The PISA-D reading literacy assessment is divided into eight proficiency levels, with the addition of Level 1c. Students at this level exhibit the following characteristics: tasks at this stage require them to understand the literal meanings of individual written words and phrases in very short sentences, grammatically simple paragraphs, or familiar contexts. Some tasks ask students to locate single words or phrases in short lists or texts based on literal matching cues. Level 1c reading texts are extremely brief, containing minimal distracting information. Texts employ familiar structures and explicit information to support student responses(OECD, 2018a) .

3.2. The Results of Reading Literacy

The results revealed urgent needs to improve students' reading literacy levels in these nations. On average, only 43% of 15-year-olds had enrolled in Grade 7 by age 15 and qualified for the PISA-D test, far lower than the 89% average in OECD countries. In Cambodia, Senegal, and Zambia, this proportion was approximately 30%.

In terms of score distribution, participating students generally showed low reading performance. Taking PISA Level 2—the baseline proficiency for "understanding simple and familiar texts literally" and demonstrating basic information integration and reasoning skills, which aligns with the minimum literacy target of SDG 4 (ensuring all youth achieve minimum proficiency in reading by 2030)—only 23% of students met this threshold, compared to 80% in OECD countries. This indicates that the average reading proficiency of 15-year-old students in these countries needs to quadruple to reach the SDG 4 target, excluding out-of-school youth (OECD, 2018b).

Country-specific data showed more severe challenges: only 5% of tested students in Zambia achieved the minimum proficiency, with an average score of 275—over two standard deviations below the OECD average of 500. Ecuador performed relatively better among participants, with 49% of students meeting the baseline and an average score of 409. Notably, since PISA-D only assesses in-school children, the actual learning levels may be lower due to upward bias in test samples—students with poorer academic performance are more likely to dropout, skewing results (OECD, 2020b).

These findings highlight the significant educational gaps in LMICs, particularly in foundational reading skills, compared to developed nations. PISA-D results provide empirical evidence for these countries to adjust education policies and optimize resource allocation, aiming to enhance students' reading literacy and overall educational quality.

4. The "Inclusivity and Equity" Contextual Questionnaire of PISA-D

4.1. The Theoretical Basis for the Construction of the Questionnaire Framework

The OECD designed the PISA-D questionnaire framework based on Willms (2018)' education success model. Willms (2018) argues that a society's ability to cultivate young people's essential and fundamental literacy depends on its capacity to provide appropriate human and material resources to support young people's healthy development from conception to adolescence. He further posits that traditional methods for measuring educational progress fail to capture the nuances of child development, and many evaluation frameworks overlook the cumulative effects of various factors influencing students' healthy development. Meanwhile, educational inputs and school-level factors are not the sole determinants of academic success; the establishment of evaluation frameworks should adopt a broader perspective to consider factors contributing to academic success (Mayer, 2009).

Willms (2018) also emphasizes that contextual information surveys in low- and middle-income countries (LMICs) should focus on indicators related to equality and equity. The "equality" pathway refers to disparities in educational outcomes across different groups. Assessments must identify relevant subpopulations, such as students of different ethnicities, those living in poverty, or students with disabilities, and accurately measure their academic performance. The "equity" pathway concerns the school resources and process conditions that influence educational outcomes and serve as the foundation for student success, including family factors like parental engagement, parent-child relationships, and parental care; institutional and school factors such as safe and inclusive environments, teaching quality, learning time, and material resources; and community capital and resources (Willms et al., 2012). Assessments should focus on disparities in access to these resources among different student groups.

4.2. Adaptation and Innovation of the PISA Questionnaire

Guided by the education success model's definition of individual academic success outcomes and foundational determinants, the PISA-D questionnaire framework comprises 15 content modules, covering four educational success outcomes, five success foundations, six equity-and-equality assessment factors, and influencing factors related to teachers, schools, and education systems. The contextual factor questionnaire for in-school assessments includes three components: student, teacher, and school surveys. For out-of-school youth assessments, PISA-D has developed youth questionnaires, parent questionnaires, and family observation questionnaires completed by specialized interviewers. In terms of measurement indicators, the PISA-D contextual factor questionnaire builds on PISA 2015 through revisions, improvements, and innovations to better measure factors more closely associated with academic performance in low- and middle-income countries (LMICs). Core indicators are retained to ensure comparability with the PISA program.

4.2.1. Indicators Related to Educational Success Outcomes

Educational success outcomes include academic performance, educational attainment, health and well-being, and attitudes toward school and learning. Academic performance in educational outcomes is measured by the reading literacy assessment, while other outcome indicators are collected through contextual factor questionnaires.

Educational Attainment, a key outcome of educational development in low- and middle-income countries (LMICs), measures the highest level of schooling students have achieved. PISA-D assesses educational attainment to understand students' pathways to their current academic performance and reasons for school dropout. Data on grade repetition, pre-school experience, attendance, chronic absence and its causes are collected through student and youth questionnaires, while information on retention policies and academic support services is obtained via school questionnaires in in-school assessments (OECD, 2018a). For out-of-school youth, the survey also focuses on employment status, weekly working hours, and wages; factors hindering youth from completing compulsory education are captured through parent questionnaires for out-of-school youth (OECD, 2018c) .

PISA defines well-being as the psychological, cognitive, social, and physical functions and capabilities necessary for students to lead happy and fulfilling lives, comprising five dimensions: cognitive well-being, psychological well-being, physical well-being, social well-being, and material well-being. Cognitive well-being, defined as the level of subject-specific skills and competencies acquired, is measured by the reading literacy assessment (OECD, 2017). Research by Helliwell et al. indicates that students in LMICs are more vulnerable to adverse factors such as hunger and chronic diseases affecting physical health (Helliwell et al., 2025). Accordingly, PISA-D asks students about their general health perceptions and experiences of anxiety and depression over the past year. Parent questionnaires for out-of-school youth collect data on prenatal and early-life experiences, including maternal health during pregnancy, childbirth complications, feeding practices within the first six months of life, and health issues before age five (OECD, 2018c). Social well-being, linked to a sense of belonging and interpersonal communication, involves students' attitudes toward school and learning. Finally, material well-being is assessed through questions on socioeconomic status, poverty, and school material resources.

Students' attitudes toward schooling and their engagement in school activities are seen as indicators of their propensity to collaborate with others and function in society, serving as critical educational outcomes for promoting lifelong learning and productive citizenship (OECD, 2003). Following PISA's framework, PISA-D measures school belonging through student questionnaires and adds questions in youth surveys about attitudes toward school, learning outcomes, and participation in learning activities. Youth questionnaires also collect information on literacy-related activities, such as frequency of reading newspapers, magazines, or books, and writing texts or emails. Parents of out-of-school youth are asked about their values and attitudes toward school education.

4.2.2. Indicators Related to Foundations for Success

Foundations for success include inclusive environments, teaching quality, learning time, instructional resources, and family and community support. In the education success model, an inclusive environment refers to a setting where all students can achieve success. "All" encompasses learners across boundaries of gender, ethnicity, nationality, religion, disability, and social class; "success" involves students' achievement in academic, physical, social, emotional, and spiritual dimensions(American University., 2019). Providing an inclusive environment is critical to educational success in LMICs, as it relates to learning opportunities for children of

diverse ethnic, linguistic, and religious backgrounds, children with disabilities, and whether these groups can enjoy learning, participate in school life, and gain full school experiences of acceptance by peers and teachers. At the institutional level, inclusivity involves policy provisions such as admission policies for diverse student groups or ability-based streaming; at the school and teacher levels, it involves whether schools provide necessary support for students with special learning needs and teachers' values toward student diversity. Specifically, PISA-D assesses school belonging, school climate, and school safety through student and youth questionnaires, and asks about school admission and grouped instruction via school questionnaires (OECD, 2019).

Effective teaching behavior is externally manifested when teachers recognize, understand, and actively pursue teaching objectives, uphold the purpose of promoting learning, and directly or indirectly transmit knowledge or content they deem valuable to students (UNESCO, 2004). PISA-D includes questions on classroom climate and teacher-student relationships in student and youth questionnaires. It also asks teachers about their attitudes and specific approaches toward teaching struggling students, and principals about teacher behaviors that may negatively impact classroom climate and teaching quality (OECD, 2019).

Addressing the issues of late school entry and high repetition rates among LMIC students, PISA-D student and youth questionnaires inquire about study time inside and outside school, reasons for absence or dropout, and causes of teacher absenteeism. Principals are asked about policies on teacher absences, specifics of reduced teaching time, and their causes (OECD, 2019).

Data from Murillo and Román (2011)' assessment of education quality in Latin America show that, even after accounting for students' socioeconomic status, school resources significantly influence academic development in LMICs. Therefore, PISA-D's measurement indicators for school material resources are more refined and comprehensive than those of PISA. PISA-D collects information on basic school services, teaching facilities, and instructional resources through school questionnaires. It also gathers principals' perceptions of resource shortages, availability of internet and ICT resources, and accessibility of teaching facilities and resources, as well as data on teachers' access to and condition of instructional resources and their usage of these resources.

Finally, family and community support reflects the social and familial dynamics that provide children with support, care, love, guidance, and protection—critical conditions for their physical and mental health development. PISA-D investigates students' and youth's communication with parents and family members through student and youth questionnaires, parents' engagement via teacher questionnaires, and how parents and communities contribute to schools through school questionnaires. It also surveys parents of out-of-school youth about the educational support they provided during their children's early adolescence.

4.2.3. Indicators Related to Equity and Equality Assessments

Grounded in Willms (2018)' education success model, which emphasizes the dual dimensions of "equality" and "equity" in educational assessment. Factors for assessing equity and equality include gender, family socioeconomic status (SES) and poverty, home language, instructional language, geographic location, migration status, and disability. PISA-D has made significant

adjustments to PISA's family SES measurement indicators by adding poverty-related metrics to more comprehensively reflect the lower educational attainment and household income levels of most students in low- and middle-income countries (LMICs). PISA-D collects data on parental highest education level, parental occupational status, and household wealth index through student and out-of-school youth questionnaires. Among these, the household wealth index is revised based on LMICs' economic development and income profiles. The questionnaire includes additional items on poverty experiences, covering material possessions, parental education, and participation in literacy activities. School questionnaires gather information on school meal programs. For out-of-school youth, questions on employment status and government education support are added. Additionally, PISA-D has specially designed family observation questionnaires to be completed by interviewers, collecting data on housing type, location, and surrounding environment to supplement SES and poverty-related information for out-of-school youth (OECD, 2018c).

4.2.4. Indicators Related to Teachers, Schools, and Education Systems

At the education system level, PISA-D collects data on national-level assessment and examination systems, teaching time, teacher training and compensation, education finance, national accounting, and population statistics. To scientifically and accurately gather information about teachers and schools related to student academic and psychosocial development in low- and middle-income countries (LMICs), PISA-D's teacher questionnaires include surveys on whether teachers teach multiple grades, engage in multidisciplinary teaching or work outside teaching, pre-service training, family socioeconomic status (SES), and health and well-being. School questionnaires add questions about school location and nearby safety conditions (OECD, 2019).

5. Discussion

The PISA-D project actively expands reading literacy assessment tools and contextual factor questionnaires, incorporates out-of-school youth assessments, and conducts corresponding analyses of related factors. Despite National Compulsory Education Quality Assessment Program in China not covering out-of-school students like PISA-D, PISA-D's experience remains valuable. Its "downward-compatible" assessment framework (e.g., Level 1c for basic literacy) can enhance our ability to identify weaknesses in disadvantaged students. The education success model-based questionnaire design, with detailed indicators on family and community support, offers insights for improving our equity monitoring systems. Furthermore, this series of development and implementation experiences of PISA-D provides important references for improving China's compulsory education Chinese subject assessment and enhancing Chinese language education and teaching.

In the context of the overall construction of the assessment system, it is crucial to draw on the experience of PISA-D to carry out specialized Chinese language subject assessment for special regions or groups and improve the national basic education quality assessment system. The Education Blue Book: Report on China's Education Development (Yang, 2019) highlights that there is a significant gap in education between urban and rural areas in China, and the

education quality in western remote provinces is relatively backward, reflecting the persistent problem of unbalanced regional education development. This situation indicates that efforts to secure baseline standards and address weaknesses in compulsory education still need to be strengthened. The academic performance of relatively disadvantaged groups, such as students in rural areas, migrant and left-behind children, children from single-parent families, as well as the teaching conditions of rural teachers, require more accurate diagnosis. China, boasting the world's largest basic education system, officially established the compulsory education quality assessment system in 2015, and the National Compulsory Education Quality Assessment Program (2021 Revised Edition) further improved this system (MOE, 2021). Nevertheless, the national basic education quality assessment system presented in the new program still resembles the main assessment in the U.S. assessment system. Given that a single assessment has limited purposes and functions, various assessment objectives should be achieved through multiple assessment projects. For example, in addition to the main assessment, the United States also conducts long - term trend assessments and specialized studies, including research focusing on specific student groups (Li et al., 2017). Therefore, China can learn from the development and implementation experience of PISA-D to further improve the national basic education quality assessment system, carrying out both regular Chinese language subject assessment and specialized assessment according to actual needs.

Regarding the design of assessment indicators and content, it is necessary to revise and improve assessment indicators and tools in a targeted manner to enhance sensitivity to special regions or groups. National-level specialized Chinese language subject assessment can draw on the construction ideas of PISA-D's reading literacy assessment tools and contextual factor questionnaires to integrate and update existing assessment indicators, enhancing the sensitivity and versatility of assessment tools, and contributing to the promotion of more equitable and high-quality education. According to the 2019 National Compulsory Education Quality Assessment Report, 18.3% of fourth-grade students and 20.7% of eighth-grade students have Chinese language academic performance at the level that needs improvement (BNU., 2022). Considering China's huge primary and secondary school student population, these proportions mean that a considerable number of students still need improvement. Following the design ideas of PISA-D, specialized assessment can adapt and update existing national basic education Chinese language subject assessment tools, conduct small-scale pre-tests, and find test items and questionnaire indicators that are more suitable for the Chinese language academic performance of disadvantaged groups and the influencing factors, providing empirical evidence and references for primary and secondary schools and teachers to formulate targeted teaching improvement measures.

6. Conclusions

In conclusion, the PISA-D represents a significant advancement in international education assessment, particularly tailored to the needs of LMICs. By expanding assessment coverage to include out-of-school youth, refining reading literacy assessment tools, and enhancing contextual factor questionnaires, PISA-D addresses the limitations of traditional PISA in capturing the

diverse educational landscapes of LMICs. Its focus on equity and inclusion, as reflected in the measurement of factors such as family socioeconomic status, poverty, and access to educational resources, offers a more comprehensive understanding of the determinants of academic success in these regions.

The multi-faceted questionnaire framework of PISA-D, informed by education success model, not only assesses educational outcomes but also delves into the foundational elements that support student development. This holistic approach provides a rich dataset for evidence-based policy-making, enabling countries to identify disparities and allocate resources more effectively. The project's "downward-compatible" proficiency levels and targeted task design further enhance its relevance, ensuring that assessments are sensitive to the varied capabilities of students in LMICs.

For China, PISA-D offers valuable insights for improving Chinese subject assessment and education. Given China's challenges of regional educational inequality and the need to support vulnerable student groups, lessons from PISA-D can guide the refinement of the national basic education quality assessment system. By incorporating specialized assessment for specific regions and populations, adapting assessment tools to measure relevant factors more accurately, and promoting equity in resource allocation, China can leverage PISA-D's experience to elevate the quality and fairness of Chinese language education. As such, PISA-D not only contributes to global educational goals but also serves as a practical model for countries seeking to enhance their domestic education assessment and policy-making processes.

Author Contributions:

Conceptualization, X.T.; methodology, X.T.; resources, X.T.; writing—original draft preparation, X.T.; writing—review and editing, X.T. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement:

Not applicable.

Informed Consent Statement:

Not applicable.

Data Availability Statement:

Not applicable.

Conflict of Interest:

The authors declare no conflict of interest.

References

American University. (2019). The Benefits of Inclusion and Diversity in the Classroom. Retrieved from Washingtion: https://soeonline.american.edu/blog/benefits-of-inclusion-and-diversity-in-the-classroom/

BNU. (2022). National Compulsory Education Quality Assessment Report. Retrieved from https://cicabeq.bnu.edu.cn/zljc/jcjgbg/c42da90750404188a74a9f844fa75d1a.htm.2018:3

Helliwell, J. F., Layard, R., Sachs, J. D., et al. (2025). World Happiness Report 2025.

Li, L., Ren, X., & Jiang, Z. (2017). Exploration and reflection on the quality monitoring of regional basic education in China. Journal of the Chinese Society of Education, 12, 37-41.

Lockheed, M., Prokic-Bruer, T., & Shadrova, A. (2015). The Experience of Middle-Income Countries Participating in PISA 2000-2015. Retrieved from Paris: https://doi.org/10.1787/9789264246195-en

Mayer, K. U. (2009). New directions in life course research. Annual Review of Sociology, 35, 413-433.

MOE. (2021). National Quality Assessment Program for Compulsory Education (Revised Edition 2021). Retrieved from http://www.moe.gov.cn/srcsite/A11/moe_1789/202109/t20210926_567095.html

Murillo, F. J., & Román, M. (2011). School infrastructure and resources do matter: Analysis of the incidence of school resources on the performance of Latin American students. School Effectiveness and School Improvement, 22(1), 29-50.

OECD. (2003). Student Engagement at School: A Sense of Belonging and Participation: Results from PISA 2000. Retrieved from Paris: https://doi.org/10.1787/9789264018938-en

OECD. (2016a). PISA 2015 Results (Volume I) : Excellence and Equity in Education. Retrieved from Paris: https://doi.org/10.1787/9789264266490-en

OECD. (2016b). PISA 2015 Results (Volume I): Excellence and Equity in Education. Retrieved from Paris: https://doi.org/10.1787/9789264266490-en

OECD. (2017). PISA 2015 Results (Volume III): Students' Well-Being Retrieved from Paris: https://doi.org/10.1787/9789264273856-en

OECD. (2018a). PISA for Development Assessment and Analytical Framework: Reading, Mathematics and Science. Retrieved from Paris: https://doi.org/10.1787/9789264305274-en

OECD. (2018b). PISA for Development Results in Focus. Retrieved from Paris: https://www.oecd.org/content/dam/oecd/en/publications/reports/2018/12/pisa-for-development_1a8aab6c/c094b186-en.pdf

OECD. (2018c). PISA for Development: Out-of-school assessment: Results in Focus", PISA in Focus. Retrieved from Paris: https://doi.org/10.1787/491fb74a-en

OECD. (2019). PISA 2018 Assessment and Analytical Framework. Retrieved from Paris: https://doi.org/10.1787/b25efab8-en

OECD. (2020a). PISA-D Out-of-School Assessment Technical Report. Retrieved from http://www.oecd.org/pisa/pisa-for-development/

OECD. (2020b). PISA for Development Out-of-school assessment Results in Focus. Retrieved from https://www.oecd.org/en/publications/pisa-for-development-out-of-school-assessment_491fb74a-en.html

UNESCO. (2004). Increasing Teacher Effectiveness (2nd ed.). Retrieved from Paris: https://unesdoc.unesco.org/ark:/48223/pf0000137629

UNESCO. (2016). Sustainable Development Goal 4 (SDG4). Retrieved from https://www.unesco.org/sdg4education2030/en/sdg4

UNESCO. (2018). The Educational Prosperity Framework: Helping Countries Provide Foundational Learning for All. Retrieved from https://uis.unesco.org/en/blog/educational-prosperity-framework-helping-countries-provide-foundational-learning-all

UNESCO. (2025). Global Education Monitoring Report. Retrieved from https://world-education-blog.org/2025/06/09/the-out-of-school-population-is-higher-than-previously-thought-and-rising/

Willms, J. D. (2018). Learning Divides: Using Data to Inform Educational Policy. Retrieved from Montreal: http://uis.unesco.org/sites/default/files/documents/ip54-learning-divides-using-data-inform-educational-policy.pdf

Willms, J. D., Tramonte, L., Duarte, J., et al. (2012). Assesing Educational Equality and Equity with Large-Scale Assessment Data: Brazil as a Case Study. Retrieved from https://www.researchgate.net/publication/254421673_Assesing_Educational_Equality_and_Equity_with_Large-Scale_Assessment_Data_Brazil_as_a_Case_Study

Yang, D. (2019). Education Blue Book: Report on China's Education Development (2019). Beijing: Social Sciences Literature Press.

iEducation, 2025, 1(2), 30-48
https://doi.org/10.71204/as5s4879

A Theoretical Framework for Interdisciplinary Instructional Design Targeting Statistical Reasoning

Yutong Wang [1], Qi Shen [1,*],

[1] Beijing Union University, Beijing 100101,China;

*** Correspondence:**

Qi Shen

sftshenqi@buu.edu.cn

Received: 27 June 2025 / Accepted: 1 September 2025 / Published online: 10 September 2025

Abstract

In the digital age, statistical reasoning (SR) is a core literacy for citizens navigating an information society. However, primary school statistics instruction has been striving to address the fragmented development of SR, disconnect between context and real life, and lack of wholeness in statistical activities. To bridge these gaps, this study proposes a novel interdisciplinary Project-Based Learning (PBL) framework grounded in Jones's four-dimensional SR model. Theoretically designed around the theme "Data-Driven Low-Carbon Action," the framework features a dual-path structure: "Longitudinal Competency Progression" (a sequenced progression through the four dimensions of Statistical Reasoning (SR), namely Describing Data Displays (D), Organizing and Reducing Data (O), Representing Data (R), and Analyzing and Interpreting Data (A)); combined with "Horizontal Disciplinary Collaboration" (integrating knowledge and practices from Mathematics, Science, Information Technology, and Chinese Language Arts). Driven by a genuine societal issue, this integrated structure aims to systematically address all dimensions of SR through a "Problem → Task → Product" spiral progression, enabling students to conceptually participate in the complete PPDAC cycle (Problem, Plan, Data, Analysis, Conclusion). This proposed framework offers a structured approach for fostering the synergistic development of students' scientific inquiry, digital innovation, language application, and SR, providing a theoretically grounded and replicable model for interdisciplinary thematic teaching design.

Keywords: Interdisciplinary Teaching; Statistical Reasoning(SR); Project-Based Learning (PBL); Primary Education

1. Introduction

With the advent of the era of big data, statistical reasoning (SR) has become an indispensable literacy for citizens to navigate in the complex information society. However, primary school instruction in statistics and probability continues to face significant challenges, including fragmented development of SR (He, 2013; Friedrich et al., 2024), disconnect between context and

real life (He, 2013; Friedrich et al., 2024), and lack of wholeness in statistical activities (Pan et al., 2022; Friedrich et al., 2024). Against this backdrop, the Mathematics Curriculum Standards for Compulsory Education (2022 ed.). (hereinafter referred to as the New Curriculum Standards) explicitly mandates that "no less than 10% of class hours in each subject be allocated to interdisciplinary theme-based learning," emphasizing the integration of multidisciplinary knowledge to address real-world problems. This policy underscores an urgent need to explore interdisciplinary instruction as a viable pathway to tackle the aforementioned challenges in primary statistics education, with the goal of systematically fostering students' SR.

Positioning this study as conceptual design research, the paper focuses on developing a theoretically sound instructional framework prior to empirical validation. Its core objective is to respond to the persistent issues in primary SR instruction by examining how interdisciplinary teaching can function as an integrative mechanism—harnessing the convergence of multidisciplinary knowledge to solve complex real-world problems. Ultimately, this research aims to provide a replicable practical paradigm for interdisciplinary thematic learning under the New Curriculum Standards, with a specific focus on enhancing students' SR.

2. Literature Review

2.1. Statistical Reasoning

2.1.1. Jones' Four-Dimensional Framework

Jones et al.(2000) established a validated framework for evaluating SR in primary school students (Grades 1-5), grounded in Biggs & Collis'(1991) cognitive development model. Through clinical interviews with 20 students, they conceptualized SR through four hierarchical dimensions:

(1) Describing Data Displays (D)

This dimension focuses on students' ability to explicitly read and communicate information directly presented in raw data or various data representations (tables, charts, graphs). It encompasses:

D1: Interpreting Data Presentation: Identifying chart elements (title, axes, labels)

D2: Recognizing Different Representations of Same Data: Distinguishing diverse forms (charts, tables) for identical data

D3: Evaluating Different Representations of Same Data: Understanding characteristics of different chart types

(2) Organizing and Reducing Data (O)

This dimension focuses on students' ability to arrange, categorize, or consolidate data into summarized forms. This ability is essential for progressing to data analysis and interpretation, as organizing data (e.g., into groups) reveals patterns and trends, and facilitates comparisons between data sets through measures of center and spread. It includes:

O1: Grouping and Sorting Data: Classifying data by logical dimensions; sorting key indicators.

O2: Awareness of Data Reduction During Reorganization: Eliminating redundant information (e.g., irrelevant fields); extracting key data

O3: Describing Data Using Terms of Typicality/Distribution: Using terms like "maximum," "proportion," "distribution" to characterize data

(3) Representing Data (R)

This dimension focuses on students' ability to display data in graphical forms. It involves developing "graphical sense," which entails understanding the construction of graphs as tools for structuring data and, crucially, selecting the most effective graphical representation for a specific context. It consists of:

R1: Completing Partially Constructed Graphs: Filling in missing data on a graph based on a given data set.

R2: Representing Different Organizations of the Same Data set: Converting data between charts, tables, text; selecting optimal representation

(4) Analyzing and Interpreting Data (A)

This dimension constitutes the core of statistical reasoning. It focuses on students' ability to recognize patterns and trends within data and to make inferences and predictions based on data. It encompasses:

A1: Identifying What the Display Does Not Reveal: Recognizing data limitations (sample, time frame); reflecting on conclusion reliability

A2: Comparing and Combining Data (Reading Between Data): Comparing associations between different datasets

A3: Inferring and Predicting from Data (Reading Beyond Data): Predicting trends based on data

2.1.2. Hierarchical Levels of Statistical Reasoning

The four-dimensional structure of SR aligns with Biggs & Collis'(1991) SOLO (Structure of the Observed Learning Outcome) taxonomy, forming a developmental continuum from low to high cognitive complexity. Jones et al. (2000) defined these levels as follows:

Level 1 (Idiosyncratic): Students focus on irrelevant information and rely heavily on subjective judgment when describing, representing, or analyzing data.

Level 2 (Transitional): Students begin using quantitative reasoning but typically focus on only one aspect of the data. Their representation and analysis remain incomplete and narrow.

Level 3 (Quantitative): Students consistently use informal quantitative reasoning and start to consider multiple aspects of the data exploration task. They tend to provide multiple responses when representing and analyzing data.

Level 4 (Analytical): Students demonstrate analytical and quantitative reasoning. They can provide multiple, logically coherent perspectives and interpretations within the data context.

2.1.3. Alignment with Curriculum Standards

The conceptualization of SR outlined above is consistent with contemporary educational goals. New Curriculum Standards emphasizes "data awareness" as a core competency, aiming to cultivate students' ability to address real-life problems by thinking through data, extracting useful information, and making reasonable inferences. The Standards designate "Statistics and Probability" as one of four key learning domains, encompassing themes such as "data classification," "data collection, organization, and expression," and "possibility of random phenomena"—all of which align closely with the four dimensions of SR defined by Jones et al.This alignment is further supported by international research. Gao (2020) noted that Jones et al.'s SR model is consistent with most international primary statistics education goals, highlighting its universal applicability.

Synthesizing the New Curriculum Standards with existing research, this study defines SR as the capacity students demonstrate in solving statistical problems to: Describing Data Displays (D), Organizing and Reducing Data (O), Representing Data (R), and Analyzing and Interpreting Data (A).

2.2. Problems in Primary SR Instruction

2.2.1. Fragmented Development of SR

The proportion of statistics and probability content in the primary school mathematics curriculum is limited, and its weight in standardized assessment is relatively low, which leads some teachers to underestimate its teaching significance (He, 2013). In traditional classrooms, instruction often mechanically focuses on reading data from charts or memorizing average formulas, and few activities are designed to cultivate SR (Friedrich et al., 2024).

2.2.2. Disconnect Between Context and Real Life

Authentic data and contexts are central to developing statistical literacy(Friedrich et al., 2024). However, there is a disconnection in primary school statistics teaching. Many teachers rely entirely on the scenarios provided in textbooks and lack the ability to transform knowledge into problem situations(He, 2013), which hinders students' understanding of the complexity and variability of real-world data.

2.2.3. Lack of Wholeness in Statistical Activities

Teachers often lack awareness of the need for complete statistical practice, frequently omitting the problem formulation stage (Pan et al., 2022). This prevents students from establishing the logical 'problem-data-decision' chain. Friedrich et al., (2024) pointed out in a systematic review that most activities are still confined to a single data processing step (e.g., graphing or calculation), making it difficult for students to understand the essence of statistics.Students often passively follow the teacher's instructions in repetitive tasks, lacking in-depth participation in the background of the problems, which stifles their interest and autonomy.

2.3. Interdisciplinary Learning

Wagner (2014) proposed that interdisciplinary thinking is "the key skill of the 21st century" and defined its essence as "the cognitive ability to identify the connections among multiple disciplines in complex problems". In STEM education, interdisciplinary studies are clearly defined as "the utilization of comprehensive knowledge and skills in science, technology, engineering and mathematics to solve real-world problems and support students' academic, professional and social success" (Halawa et al., 2024). Focusing on school settings, Xia (2022) further defined Interdisciplinary Project-Based Learning (Interdisciplinary PBL), emphasizing that students, to solve a complex and authentic driving question, need to engage in creative knowledge integration, ultimately forming integrated understanding and products that transcend single disciplines, rather than mere juxtaposition of subject knowledge. Guo and Yuan (2023) categorized interdisciplinary thematic learning along two dimensions: knowledge status and disciplinary relationships. Regarding knowledge status, it can be "applying knowledge to solve complex problems" or "learning knowledge through interdisciplinary themes." Regarding disciplinary relationships, it can be either "Single Discipline-Led" or "Multiple Disciplines-Led."

Synthesizing these perspectives, this study defines interdisciplinary learning as: An interdisciplinary thematic learning process anchored in a genuine and complex real-world problem, employing a "Single Discipline-Led" model, which guides students to apply and integrate multi-disciplinary knowledge, methods, and skill to solve the core problem.

2.4. Interdisciplinary Teaching Facilitates SR Development

Traditional primary statistics instruction faces core problems—fragmented competency development, lack of authentic contexts, and incomplete activities—in cultivating SR. This study argues that Interdisciplinary Teaching, particularly through Interdisciplinary PBL, provides an effective pathway to address these issues and systematically foster SR.

2.4.1. Constructs a Systemic Competency Framework, Addressing Disjointed Competency Development

SR development involves a coherent cognitive process spanning four dimensions: Describing Data Displays (D), Organizing and Reducing Data (O), Representing Data (R), and Analyzing and Interpreting Data (A) (Jones et al., 2000) However, traditional classrooms often divide statistical knowledge into isolated skills such as chart reading or formula calculation, hindering students' grasp of the holistic logic of SR.

This division stems from the neglect of the need to integrate the cognitive, behavioral and affective dimensions in the development of statistical literacy (Friedrich et al., 2024). Watson et al.(2003) suggested mathematics can be integrated with other subjects to provide social contexts for statistical learning and foster decision-making skills. Interdisciplinary teaching, by utilizing real problems to integrate statistics with practices from science and technology, has the potential to form a coherent competency development path that addresses such fragmentation.

2.4.2. Leverages Authentic Societal Issues, Realizing Data Value Transformation

Real data and background are the core of the development of statistical literacy (Friedrich et al., 2024). However, teachers often rely solely on textbook data, neglecting the motivating power of real problems (Liu, 2020). This gap necessitates context reconstruction. Interdisciplinary teaching, oriented towards socio-economic realities (Gibbons, 1994) bridges the classroom-reality divide through its contextualized knowledge production mode. The research by Vahey et al.'s (2012) further indicates that addressing genuine interdisciplinary issues such as "water resource usage" or "climate change" can significantly enhance students' key data interpretation skills and their willingness to apply what they have learned to life.

2.4.3. Covers the Complete Statistical Inquiry Cycle, Preventing Activity Discontinuity

Traditional instruction often limits statistical activities to single steps (e.g., graphing), resulting in passive execution and a lack of experience with the entire process. When embedded in authentic measurement contexts, the proportion of students completing the full cycle (design-collect-represent-interpret) increases significantly: 58% of the students could independently plot and interpret the data distribution, and 39% of the students could compare the distribution in different scenarios (English & Watson, 2015). Combining this with Xia 's (2022) progressive interdisciplinary prototype—emphasizing decomposing sub-problems according to problem-solving logic to form a complete inquiry chain with "dynamic disciplinary synergy and progressive competency development" —avoids fragmentation. Such an approach ensures that statistical activities are not isolated but part of a continuous, purposeful inquiry process.

In summary, interdisciplinary teaching offers tripartite support for developing Statistical Reasoning (SR) through its systemic framework design, authentic problem contexts, and complete inquiry cycles. However, current practices have yet to systematically integrate Jones' cognitive dimensions (D-O-R-A) with disciplinary collaboration mechanisms. The dual-dimensional framework ('Longitudinal Competency Progression + Horizontal Disciplinary Collaboration') proposed in this study responds to this limitation, seeking to enable synergistic development across SR's four dimensions via a structured pathway.

3. "Data-Driven Low-Carbon Action" Project Design

3.1. Project Overview

Project Title: Data-Driven Low-Carbon Action

Driving Question: How can we use data to persuade community residents to participate in low-carbon action?

Core Competency Goals: Develop SR; enhance low-carbon awareness and agency.

Core Product Chain: Greenhouse Effect Lab Report (Math/Science) → Household Carbon

Footprint Analysis Report (Math/Science) → AI Creative Low-Carbon Poster (Math/IT) → Letter to Community Residents (Math/Chinese).

3.2. Interdisciplinary Integration Framework

The project established a two-dimensional "Longitudinal Competency Progression + Horizontal Disciplinary Collaboration" framework (Figure 1). Longitudinally, it uses the D-O-R-A model (Describing Data Displays → Organizing and Reducing Data → Representing Data → Analyzing and Interpreting Data) as the SR development path. Horizontally, it establishes a Mathematics-led disciplinary network: Mathematics provides core statistical knowledge/methods; Science provides anchoring contexts (e.g., greenhouse experiment) and principle support (e.g., carbon neutrality); Information Technology (IT) enables tool application (e.g., AI poster design); Chinese Language Arts drives product output and persuasive logic (e.g., letter argumentation). Disciplines dynamically align with lesson competency goals, forming a progressive product chain ("Lab Report → Carbon Footprint Report → AI Poster → Advocacy Letter") serving the driving question: "Persuading the community with data."

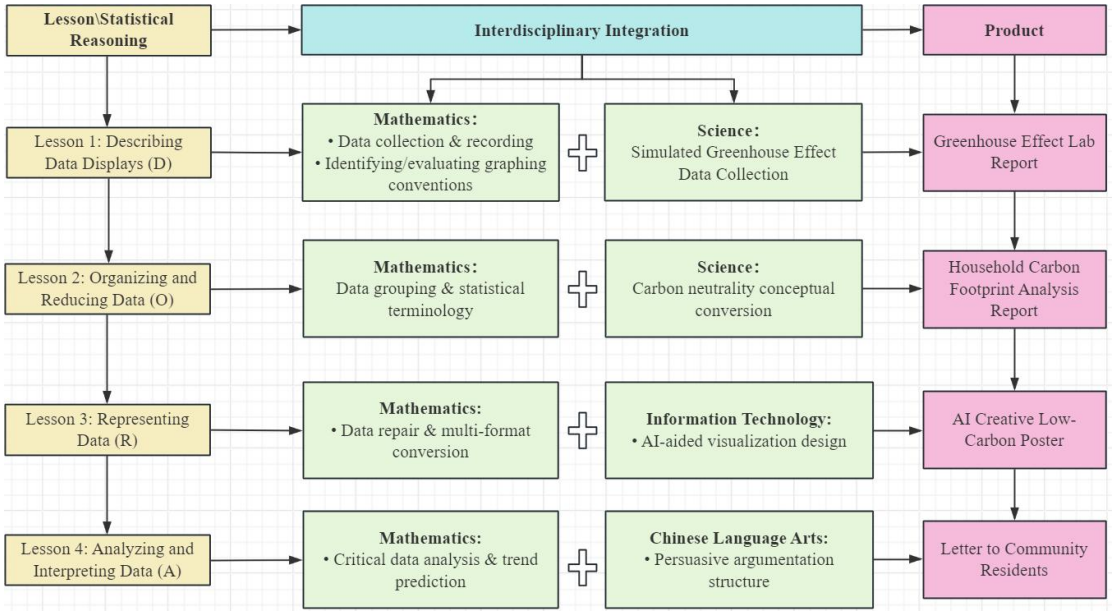

Figure 1. Interdisciplinary Integration Framework

3.3. Project Activity Flow

Centered on the driving question "How can we use data to persuade community residents to participate in low-carbon action?", the project designs four interdisciplinary lessons targeting the four SR dimensions (D, O, R, A), achieving a transformation from "Scientific Experiment Data → Household Lifestyle Data → Community Action Data." The specific flow is shown in Figure 2.

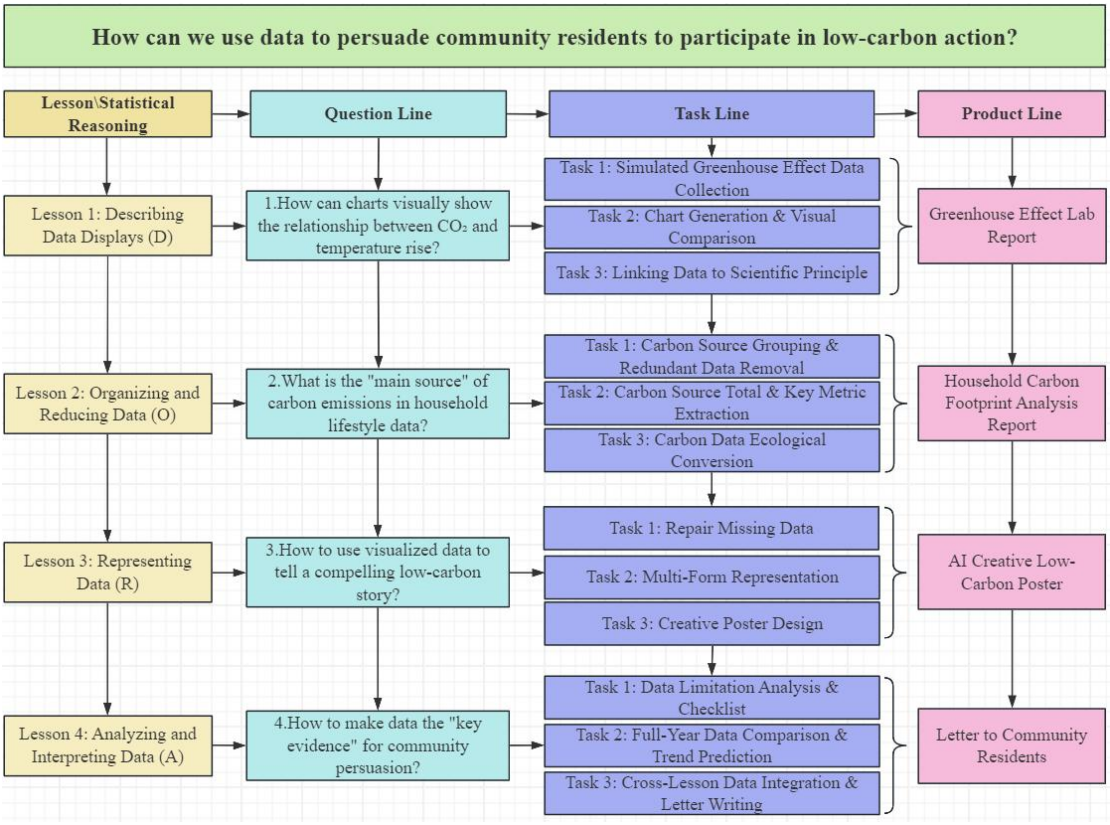

Figure 2. Project Activity Flow

3.4. Lesson Process Design

Lesson 1: Describing Data Displays (D)

Sub-question: How can charts visually show the relationship between CO_2 and temperature rise?

Product: Greenhouse Effect Lab Report

Task 1: Simulated Greenhouse Effect Data Collection

(1) Conduct sealed bag + dual thermometer experiment. Record internal/external temperature (°C) every minute (5 min).

(2) Calculate temperature difference. Record in Temperature Change Log, labeling units (°C).

(3) Competency Goal: Master basic data collection and recording skills (D1 Prep).

Task 2: Chart Generation & Visual Comparison

(1) Observe bar chart and line chart corresponding to log (X-axis: Time/min; Y-axis: Temperature/°C). Compare internal/external temperature trends.

(2) Group discussion & recording:

a. How do bar and line charts present data differently?

b. Which chart more intuitively warns about CO_2 concentration and climate crisis linkage? Why?

Competency Goal: Interpret chart elements (D1); Recognize presentation features of different charts (D2); Evaluate chart effectiveness (D3).

Task 3: Linking Data to Scientific Principle

Explain "Why does the sealed bag heat up more?" using data, linking to CO_2 greenhouse effect principle.

Competency Goal: Derive scientific conclusions from data phenomena (D1 Deep Interpretation).

Lesson 2: Organizing and Reducing Data (O)

Sub-question: What is the "main source" of carbon emissions in household lifestyle data?

Product: Household Carbon Footprint Analysis Report

Task 1: Carbon Source Grouping & Redundant Data Removal

(1) Classify household weekly carbon emission data (containing redundant fields like "usage period," "device ID") into: Appliances / Transport / Gas / Waste / Other.

(2) Mark irrelevant columns for deletion (e.g., "device ID"), justifying removal (e.g., "unrelated to source type").

Competency Goal: Group data by logical dimensions (O1); Identify invalid information (O2).

Task 2: Carbon Source Total & Key Metric Extraction

(1) Calculate weekly totals for 5 source types. Record in simplified table ("Source Type," "Carbon Emission").

(2) Describe data distribution using statistical terms:

a. Which source has the highest emissions? What proportion?

b. What is the average daily household carbon emission?

Competency Goal: Use terms like "maximum," "proportion" to describe distribution (O3).

Task 3: Carbon Data Ecological Conversion

(1) Annual Emission Estimate: Annual Emission (kg) = Weekly Total × 52.

(2) Carbon Neutrality Concept: Briefly explain: Achieving net-zero carbon emissions via reduction and carbon absorption (e.g., afforestation). Emphasize "reduction is core, absorption is supplementary."

(3) Afforestation Offset Estimate: Trees Needed = Annual Household Emission (kg) ÷ Average Annual Carbon Sequestration per Tree (kg/tree). (Note: Teacher provides sequestration data; focus is on understanding the offset concept).

Competency Goal: Establish quantitative link between data and real-world ecology.

Lesson 3: Representing Data (R)

Sub-question: How to use visualized data to tell a compelling low-carbon story?

Product: AI Creative Low-Carbon Poster

Task 1: Repair Missing Data

Complete missing data on a bar chart of weekly average AC carbon emissions for a household.

Competency Goal: Understand impact of data completeness on charts (R1).

Task 2: Multi-Form Representation

(1) Convert repaired bar chart into: (a) Statistical table (daily emissions); (b) Line chart (trend arrows). Compare forms: Which best shows "AC carbon emission trend"? Why?

(2) Competency Goal: Master multiple representations of same data (R2).

Task 3: Creative Poster Design

(1) Extract core data insights (e.g., "Weekend peak," "Summer trend"). Combine with "Smart AC Use" concept. Input into AI tool to generate poster.

(2) Poster must include:

a. ≥1 statistical chart (bar/line).

b. Low-carbon message

c. Color elements (e.g., green theme for eco-friendliness).

Competency Goal: Use visualization to convey data narrative (R2 Application).

Lesson 4: Analyzing and Interpreting Data (A)

Sub-question: How to make data the "key evidence" for community persuasion?

Product: Letter to Community Residents

Task 1: Data Limitation Analysis & Checklist

(1) Discuss limitations of previous lesson's “Single-Week AC Carbon Emission Line Chart”:

Time: Covers seasonal (winter/summer) differences?

Sample: Can single household represent whole community?

(2) Complete Data Blind Spot Checklist: Record ≥2 limitations and their impact on conclusions.

Competency Goal: Critically analyze data boundaries (A1).

Task 2: Full-Year Data Comparison & Trend Prediction

(1) Analyze household's 12-month AC carbon emission line chart:

Peak Month: _____, Reason _____ (e.g., "8h daily use in heat").

Secondary Peak: _____, Reason _____ (e.g., "heating + AC").

(2) Calculate annual total. Assume 2 % annual growth, predict next year's total.

Competency Goal: Compare/combine multi-source data (A2); Infer trends from data (A3).

Task 3: Cross-Lesson Data Integration & Letter Writing

(1) Select ≥2 data types from previous lessons (e.g., Lesson 1 temp diff, Lesson 3 weekend peak).

(2) Write letter using Letter Framework:

a. Science Intro: Cite Lesson 1 experiment on CO_2-emissions link.

b. Data Warning: Use Lesson 2 "Appliances = 58% emissions" & Lesson 4 "Summer peak XXkg" to argue problem severity.

c. Action Proposal: Based on Lesson 3 "Raise thermostat 1°C saves 19kg/month," propose collective community reduction goal.

Competency Goal: Build persuasive logic with data chain (A2, A3); Apply formal letter format (Chinese integration).

3.5. Product Showcase & Assessment

During the showcase, groups sequentially present their interconnected products guided by "Data-Driven Low-Carbon Action": (1) Present Greenhouse Effect Lab Report via visualizer, explaining CO_2-temperature relationship using charts. (2) Explain process of organizing/reducing household carbon data using the Carbon Footprint Analysis Report; Present source totals and analysis (max, proportion) via statistical table; Explain carbon neutrality and afforestation calculation. (3) Showcase AI Low-Carbon Poster, explaining chart choice, message, and color design intent. (4) Read the Letter to Community Residents.

An interactive Q&A follows, where other groups/teachers pose questions. Presenting groups respond, clarifying challenges and highlights. Finally, using the assessment rubric (see Appendix A) students conduct self-assessment, peer assessment (within group), and teacher assessment based on SR, Interdisciplinary Literacy, and Collaboration, comprehensively evaluating group performance and learning outcomes, guiding deep reflection to enhance SR and interdisciplinary literacy.

4. Discussion

4.1. Innovations and Theoretical Contributions

This study addresses three core issues in primary statistics teaching—fragmented competency development, lack of authentic contexts, and incomplete activities—by constructing a dual-dimensional framework ("Longitudinal Competency Progression + Horizontal Disciplinary Collaboration"). This framework embeds Jones' (2000) four-dimensional SR model into the interdisciplinary PBL structure. Distinct from traditional single-subject designs, this Mathematics-led framework dynamically integrates Science (experimental context), IT (tool empowerment), and Chinese Language Arts(argumentation/expression), enabling systematic SR development within authentic problem-solving. This responds to Friedrich et al.'s (2024) call for interdisciplinary data literacy cultivation and conceptually aligns with the effectiveness of Xia 's (2022) "progressive interdisciplinary prototype" in statistics education.

Most importantly, this project transforms abstract statistical concepts into concrete cognitive experiences through the social science topic of "Data-Driven Low-Carbon Action." For instance, in the "Household Carbon Footprint Analysis" task, students' quantitative operations (calculating emissions, trees for offset) not only build mastery of terms like "distribution" and "proportion" (O3) but also deepen understanding of data's value as a tool for environmental decision-making (A2, A3). This is consistent with the assertion of Watson et al.'s (2003) that combining statistical data with real-world problems (such as climate change) can cultivate critical data interpretation skills.

4.2. Practical Implications for SR Development

This research provides a replicable and practical model for the teaching reform in the field of statistics at the primary school stage.

By designing the interdisciplinary project "Data-Driven Low-Carbon Action", it demonstrates how the integration of multidisciplinary knowledge can enhance SR. This model can be widely applied to other domains, supporting the cultivation of talent with interdisciplinary literacy and innovative capacity.

Furthermore, the study emphasizes the significance of real data and real-world contexts in statistics teaching. By guiding students to collect, analyze, and interpret real data, it helps them establish connections between data and the real world, enhancing data application and problem-solving skills.

4.3. Construction of a Support System for Teachers' Interdisciplinary Competency

The successful implementation of Interdisciplinary Project-Based Learning (such as the "Data-Driven Low-Carbon Action" project) imposes extremely high requirements on teachers' interdisciplinary literacy. This not only includes teachers' profound content knowledge in their own discipline (e.g., Mathematics) but also requires sufficient understanding of collaborative disciplines (Science, Information Technology, Chinese Language Arts) to effectively integrate their concepts, methods, and practices. To address this critical challenge, it is necessary to provide teachers with a systematic and practical support system. To enhance the practical feasibility of the framework and empower teachers, the following support strategies are proposed:

4.3.1. Hierarchical Training System

Stage 1: Foundation and Philosophy

Introduce the theoretical basis for cultivating interdisciplinary statistical reasoning, the D-O-R-A model, principles of Project-Based Learning, and disciplinary integration models (e.g., the Mathematics-led model).

Stage 2: Collaborative Design and Planning

Provide tools and processes for teachers of different disciplines to collaboratively design project units, including identifying driving questions, mapping the statistical reasoning development pathway to activities, clarifying the specific contributions and integration points of each discipline, planning assessment schemes, and allocating resources.

Stage 3: Implementation and Guidance

Focus on classroom management strategies for interdisciplinary projects, guiding students in interdisciplinary inquiry, integrating technologies (e.g., AI tools for visualization), assessing interdisciplinary outcomes, and facilitating effective group collaboration.

4.3.2. Typical Case Studies and Lesson Repository

Organize and develop detailed cases of successfully implemented interdisciplinary statistical reasoning projects, particularly those framed by the D-O-R-A development pathway. The "Data-Driven Low-Carbon Action" project itself should be documented as a core case. These cases should include: detailed lesson plans highlighting integration points and key focuses of statistical reasoning dimensions for each lesson; samples of students' works and final outputs; teachers' reflections on challenges encountered and corresponding solutions; video or textual records of key classroom interactions. Exposure to such concrete cases can provide teachers with valuable examples for adaptation.

4.3.3. Construction of an Interdisciplinary Teaching and Research Community

It is recommended that schools establish fixed teams consisting of teachers from Mathematics, Science, Information Technology, Chinese Language Arts, and other disciplines, equipped with curriculum coordinators or subject experts to provide professional guidance, thereby forming a supportive community characterized by "teacher collaboration + expert leadership." This community should focus on authentic interdisciplinary themes (e.g., "Data-Driven Low-Carbon Action"), clarify the core contributions and integration nodes of each discipline in tasks through regular joint lesson preparation (e.g., data analysis in Mathematics, principle support in Science, expression logic in Chinese Language Arts), and avoid the mere superimposition of disciplinary knowledge. It should also conduct cross-classroom observations and mutual evaluations to refine strategies for disciplinary integration.

4.4. Future Research Directions

Future research should first focus on empirical verification of the proposed framework. As a conceptual design, the effectiveness of the framework in cultivating students' statistical reasoning (SR) requires systematic empirical testing. Subsequent studies could adopt mixed research methods, such as pre-post tests to measure changes in students' SR competencies (using the assessment rubric in Appendix A), classroom observations to record implementation processes, and interviews with students and teachers to collect data on learning outcomes and practical implementation barriers. Such empirical evidence will help refine the framework and enhance its practical applicability.

Deepening Teacher Professional Development Support: As highlighted in Section 4.3., teacher interdisciplinary literacy remains a critical challenge and a key area for future research.

Exploring Broader Applications: Future research should further explore the potential of interdisciplinary teaching in other subject areas. Designing and implementing more innovative and practical projects, integrating knowledge and methods from different disciplines, is of great significance for cultivating compound talents.

Appendix A: Project Assessment Rubric

Primary Dimension	Secondary Dimension	Key Elements	Scoring Standard
Statistical Reasoning	Describing data displays (D)	D1: Interpret elements of experimental data charts/graphs (e.g., title, axes, labels).	1: Fails to identify any chart elements 2: Identifies isolated elements without data relevance 3: Identifies partial core elements (e.g., title/values) with incomplete descriptions 4: Can fully identify and explain all chart elements (title, horizontal/vertical axes, values)
		D2: Recognizing Different Representations of Same Data	1: Unable to distinguish chart types or provides irrational justifications 2: Distinguishes types based on superficial features (e.g., "look different") 3: Correctly identifies types but lacks analytical justification 4: Correctly distinguish chart types and interpret them by combining with the data in the charts.
		D3:Evaluating Different Representations of Same Data	1: Irrelevant reasons based solely on subjective preferences (e.g., "Bright colors") 2: Evaluation based on non-data factors (e.g., "Line charts are faster to draw") 3: Points out a single advantage (e.g., "Line charts can show changes") 4: Conducts multi-dimensional evaluation (e.g., "Bar charts are suitable for comparing instantaneous values"; "Line charts are suitable for showing CO_2-induced temperature rise trends"; "The slope angle shows the acceleration of temperature rise")
	Organizing and reducing data (O)	O1: Group household carbon emission data by source type.	1: No grouping or irrelevant grouping criteria 2: Chaotic grouping (e.g., classifying air conditioners separately) 3: Correctly completes single-dimensional grouping (e.g., all electrical appliances are classified as "home appliances" without further subdivision) 4: Groups carbon emission sources by multiple dimensions (e.g., "Classifies air conditioners as electrical appliances and gasoline-powered vehicles as transportation") and explains the

			grouping logic
		O2: Identify and remove redundant information from the data table.	1: Does not delete any information or deletes data randomly 2: Can delete some redundant information but retains a small amount of irrelevant fields 3: Can delete major redundant information and briefly explain the reason (e.g., "Unrelated to carbon emissions") 4: Can comprehensively identify redundant fields (e.g., "device ID", "usage period") and explain the reason for deletion and its relevance to data
		O3: Use statistical terms like "maximum" and "proportion" to describe the distribution characteristics of carbon emissions.	1: Unable to use statistical terms to describe data. 2: Can use non-precise terms such as "most" and "half" to describe distribution 3: Can accurately point out the source of the maximum value (e.g., "Air conditioners have the highest emissions") and calculate simple proportions 4: Can calculate proportions based on data (e.g., "Transportation emissions account for 35%") and explain their significance for low-carbon actions
	Representing data (R)	R1: Understand and complete missing data in charts/graphs.	1: Fills in randomly or leaves blank 2: Completes the conversion but omits key information such as trend arrows 3: All data is reasonable but no method is explained 4: Accurately completes based on data patterns and explains the method
		R2: Convert air conditioner (AC) carbon emission data between representations (e.g., bar chart, statistical table, line chart).	1: Unable to convert or with incorrect formats 2: Converts formats but omits key information such as trend arrows 3: Correctly completes chart/table conversion but with brief reasons 4: Correctly completes chart/table conversion and explains advantages (e.g., "Line charts can highlight the changing trend of air conditioner carbon emissions")
	Analyzing and interpreting	A1: Critically analyze the limitations of a	1: Fails to identify limitations 2: Can simply mention incomplete data but does not explain the impact

	data (A)	single data source.	3: Can point out one key limitation (e.g., "Data only covers 1 week") 4: Can, in combination with research purposes (e.g., community promotion), point out the sample limitations of the data (e.g., "Data from only 1 household") and time limitations (e.g., "Does not include winter")
		A2: Compare and combine data across charts/graphs.	1: Unable to identify "Peak Month" or "Secondary Peak", or confuses data (e.g., misjudges the lowest month as the highest month), with reasons unrelated to data (e.g., "Guessed randomly") 2: Can correctly identify either "Peak Month" or "Secondary Peak" but describes reasons without referring to data 3: Can accurately identify both "Peak Month" and "Secondary Peak" and describes reasons with reference to data 4: Can accurately identify both "Peak Month" and "Secondary Peak" and describes reasons based on data patterns
		A3: Make simple trend predictions for carbon emissions based on existing data.	1: Unable to make any predictions 2: Can simply guess the trend (e.g., "It will keep increasing") 3: Can accurately calculate annual total emissions (e.g., the sum of 12-month data is 1200kg) and correctly predict the total for next year with a 2% growth (1200×1.02=1224kg) but does not explain the prediction basis (e.g., fails to state "The 2% growth is based on the assumption that household electricity usage habits are stable") 4: Can accurately calculate annual total emissions, correctly predict the total for next year with a 2% growth, and explain the prediction logic (e.g., "Household electricity usage habits have been stable in the past 3 years, so a 2% growth is assumed; if energy-saving measures are promoted, the growth may drop to 0.5%"), reflecting consideration of data limitations
Interdisciplinary Literacy	Scientific Inquiry Practice	1. Derive scientific conclusions (greenhouse	1: Only repeats experimental steps without mentioning data and principles 2: Can describe experimental phenomena (e.g., "The temperature inside the bag is higher") but

		effect principle) from experimental data phenomena.	fails to link to principles 3: Can link data to the greenhouse effect and explain "More CO_2 → Higher temperature" 4: Can comprehensively explain the heat preservation effect of CO_2 based on temperature difference data (e.g., "3℃ higher inside the sealed bag")
		2. Understand the concept of carbon neutrality and correctly calculate the number of trees needed to offset annual household carbon emissions.	1: Unable to complete calculations or does not understand the meaning of formulas 2: Has minor calculation errors but understands the offset concept 3: Can correctly calculate results and briefly explain the meaning of carbon neutrality 4: Can accurately calculate annual emissions and the number of trees needed, and explain the logic that "Emission reduction takes priority over offset"
	Digital Innovation	1.Poster Completeness: Clearly tells a low-carbon story by integrating at least one statistical chart, text message, and visual image elements.	1: Missing elements (e.g., no charts) with vague information 2: Contains 1 element with complete information but lacking relevance 3: Contains 2 elements with close correlation between charts and text 4: Contains 3 or more elements (e.g., line chart + slogan + icon) with coherent logic (e.g., "Trend chart → Problem → Action")
		2.Creative Expression: Uses color/layout and other visual elements to enhance the persuasive power of the data.	1: Chaotic colors with obscured data 2: Single color with basically reasonable layout 3: Colors conform to the theme, layout is clear, and data is placed reasonably 4: Uses green tones to symbolize environmental protection, with layout highlighting core data (e.g., enlarging peak values) to enhance visual impact
	Language Application	1. Richness of data sources: Integrates data from at least	1: Does not use any data 2: Integrates data from 1-2 lessons but with complete irrelevance (e.g., using "experimental date" from Lesson 1 to discuss "carbon

		two project lessons.	emission issues") 3: Integrates data from 2 lessons and clearly distinguishes data types 4: Integrates data from 3-4 lessons (e.g., experiment + household + community), marks sources (e.g., "Table 1-1 experimental data"), and proposes suggestions with a data chain
		2. The community letter follows standard formal letter format.	1: Chaotic format (e.g., confusing salutation and body text) 2: Complete format but missing key parts (e.g., no signature) 3: Basically complete format 4: Includes salutation, body text (data argumentation), signature, and date, with standard format and appropriate tone
Collaborative Practice	Teamwork	1. Clear division of labor within the group with reasonable task allocation.	1: No division of labor or chaotic division of labor 2: Has division of labor but some members have vague tasks 3: Reasonable division of labor with clear tasks for each member 4: Fixed roles (e.g., recorder, analyst) with tasks matching abilities
		2. Mutual help and communication among group members to jointly solve problems encountered during the project.	1: Only completes personal tasks without participating in collaboration 2: Can participate in discussions but with few contributions 3: Can cooperate to solve problems and provide 1-2 effective suggestions 4: Takes the initiative to help peers and solves 3+ problems through discussions
	Project Output	Completed all project tasks on time and produced high-quality deliverables.	1: Unfinished with many errors 2: Basically completes tasks with a small number of issues needing revision 3: Completes all tasks on time with error-free outcomes 4: Completes all tasks in advance with outcomes exceeding expectations

Scoring Levels: 1 point: Idiosyncratic; 2 points: Transitional; 3 points: Quantitative; 4 points: Analytical.

Author Contributions:

Conceptualization, **Yutong Wang**; methodology, **Yutong Wang;** writing—original draft preparation, **Yutong Wang**; writing—review and editing, **Qi Shen**; supervision, **Qi Shen**. All authors have read and agreed to the published version of the manuscript.

Funding:

This research received no external funding.

Institutional Review Board Statement:

Not applicable.

Informed Consent Statement:

Not applicable.

Data Availability Statement:

Not applicable.

Acknowledgments:

The author expresses profound gratitude to Professor Qi Shen for his expert guidance and critical advice throughout this research. His rigorous scholarship and intellectual mentorship were fundamental to the completion of this work.

Conflict of Interest:

The authors declare no conflict of interest.

References

Biggs, J. B., & Collis, K. F. (1991). Multimodal learning and intelligent behavior. In H. Rowe (Ed.), Intelligence: Reconceptualization and measurement (pp. 57–76). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

English, L. D., & Watson, J. M. (2015). Exploring variation in measurement as a foundation for statistical thinking in the elementary school. International Journal of STEM Education, 2(1), 3.

Friedrich, A., Schreiter, S., Vogel, M., Becker-Genschow, S., Brünken, R., Kuhn, J., ... & Malone, S. (2024). What shapes statistical and data literacy research in K-12 STEM education? A systematic review of metrics and instructional strategies. International Journal of STEM Education, 11(1), 58.

Gao, X. (2020). The Construction of Teaching Mode Focusing on the Cultivation of Statistical Reasoning for Elementary School Students——Take Grade 6 students in S school as an example [Doctoral dissertation, East China Normal University].https://link.cnki.net/doi/10.27149/d.cnki.ghdsu.2020.000191doi:10.27149/d.cnki.ghdsu.2020.000191.

Gibbons, M., Limoges, C., Scott, P., Schwartzman, S., & Nowotny, H. (1994). The new production of knowledge: The dynamics of science and research in contemporary societies. Sage Publications.

Guo, H., & Yuan, Y. (2023). Basic types and implementation key points of interdisciplinary thematic learning. Primary and Middle School Management, (5), 10–13.

He, W. (2013). Research on Cognitive Development of Statistical in12–15 Years Old [Doctoral dissertation, Hangzhou Normal University].https://kns.cnki.net/kcms2/article/abstract?v=FqAfUZ3F7bafeNF3PYhkI4bQ3Y3-w7f89Gcdw224dQ7xA2wfZJhydQe_FyZ68n2-P7nrYCB-cyZMdwbhSFsPhBTbKlgQFksY8KjBYIRzDc93s0r7Iy7uzl8WMudUJBc7Ag5GZIjxBa-dWKrn6OEiJOXQcijaVa3xiehoOHwIaXX5KavmZ-f5mg==&uniplatform=NZKPT&language=CHS

Halawa, S., Lin, T. C., & Hsu, Y. S. (2024). Exploring instructional design in K-12 STEM education: a systematic literature review. International Journal of STEM Education, 11(1), 43.

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.

Liu, X. (2020). The Value and Its Realization Path of Primary-School Statistical Education.Theory and Practice of Education, 40(32), 62-64.

Ministry of Education of the People’s Republic of China. (2022). Mathematics curriculum standards for compulsory education (2022 ed.). Beijing Normal University Press.

Pan, Y., Li, Y., Xu, W., Liu, W., Liu B., & Chen, X. (2022). Topics,Trends and Enlightenment of International Research in Statistics Education——Review of International Handbook of Research in Statistics Education. Journal of Mathematics Education, 31(5), 82–89.

Vahey, P. , Rafanan, K. , Patton, C. , Swan, K. , Mark van ’t Hooft, & Kratcoski, A. , et al. (2012). A cross-disciplinary approach to teaching data literacy and proportionality. Educational Studies in Mathematics, 81(2), 179-205.

Wagner, T. (2010). The global achievement gap: Why even our best schools don't teach the new survival skills our children need-and what we can do about it. ReadHowYouWant. Com.

Watson, J., & Callingham, R. (2003). Statistical literacy: A complex hierarchical construct. Statistics Education Research Journal, 2(2), 3-46.

Xia, X. (2022). Inter-subject Project-based Learning:Definition,Design Logic,and Practical Prototype. Curriculum, Teaching Material and Method, 42(10), 78–84.

iEducation, 2025, 1(2), 49-60
https://doi.org/10.71204/twtqbt77

A Comparative Study of Middle School Mathematics Textbooks in China and Russia

Hongzhu Wang[1]

[1] School of Mathematics and Statistics, Yancheng Teachers University, Yancheng 224000, China

***Correspondence:**

Hongzhu Wang

wanghz@yctu.edu.cn

Received: 14July 2025/ Accepted: 8 September 2025 / Published online: 10 September 2025

Abstract

This study compares and analyzes the presentation of rational number concepts in two versions of junior high school mathematics textbooks from China and Russia. The findings reveal that while both versions cover core concepts consistently, there are significant differences in their definitional logic and organizational sequence. These differences stem from distinct editorial philosophies: the Russian version emphasizes the intrinsic logic and structural coherence of number system expansion, whereas the Chinese version prioritizes practical knowledge application and cognitive continuity. Based on these findings, the study proposes restructuring the conceptual sequence, clarifying functional positioning, and enhancing the rigor of definitions. It advocates that textbook development should balance logical depth with cognitive accessibility, providing empirical references for cross-cultural mathematics textbook design.

Keywords: Textbook Comparison; Rational Numbers; Extension of Number Fields

1. Introduction

Mathematics textbooks serve as the core medium for knowledge transmission, and their design philosophy profoundly influences students' cognitive structures and logical thinking development. In the era of globalized education, cross-national textbook comparison studies have become an important approach for revealing cultural differences and optimizing curriculum design (Fan et al., 2013). The presentation of foundational mathematical concepts such as number system extensions directly impacts the development of students' algebraic thinking (Van Hoof et al., 2017), while also embodying rich historical content. Consequently, number system extension has become a focal point in comparative textbook research. Internationally, studies on this core concept have long centered on comparisons between Chinese and American, or Chinese and Japanese, textbooks: American materials tend toward “contextualized applications” (such as mathematical

modeling emphasized in PISA assessments), while Japanese textbooks excel in "meticulously designed cognitive progression" (Shimizu, 2009).

However, compared to other cross-national comparisons, comparative research on Chinese and Russian mathematics textbooks remains relatively limited and has yet to form a systematic research framework. Russian mathematics education has profoundly influenced other countries, particularly socialist nations (Alexander and Vogeli, 2010). Since the 1950s, Soviet educational influence on China has been comprehensive and profound. Russian mathematics education is renowned for its axiomatic systems and logical rigor (Karp, 2007), with textbook compilation strictly adhering to the "spiral progression" principle (e.g., Kolmogorov's Introduction to Algebra and Analysis systematically introduces analytical concepts); In contrast, Chinese textbooks emphasize real-world connections and cognitive adaptability (Cai and Jiang, 2017), prioritizing "progressive understanding" through everyday examples (e.g., the Jiangsu Education Press textbook introduces negative numbers via scenarios like temperature and debt). This divergence stems from the deep tension between "formal rigor" and "intuitive perception" in mathematics education (Sfard, 2007), rooted in each nation's educational traditions: Russia inherits Vygotsky's "theory-first" principle (emphasizing top-down transmission of scientific concepts) (Vygotskij, 1979), while China remains profoundly influenced by the Confucian educational philosophy of "gradual progression" (e.g., the cognitive path of "reviewing the old to understand the new" in The Analects) (Li, 2004). The collision of these two approaches offers a unique perspective for mathematics education design.

Existing research indicates fundamental differences in how expanded number systems are presented in textbooks: one approach emphasizes cognitive pathways that reduce students' cognitive load (e.g., making algebraic definitions more intuitive), while the other advocates for structured mathematical logic (Wu, 2011). The representative textbooks selected for this study—China's Su Jiao Edition Mathematics Grade 7 Volume 1 (2024) and Russia's PR Edition Mathematics Grade 6 Volume 2 (2023)—precisely embody these contrasting approaches. Therefore, this paper focuses on the conceptual construction process of rational numbers, conducting a comparative analysis of the relevant conceptual systems. This aims to yield valuable educational insights.

2. Research Subjects

2.1. Selection of Textbooks

The textbooks published by Jiangsu Phoenix Science and Technology Press are widely adopted in one of China's most economically and educationally developed provinces. They embody the principles of China's Compulsory Education Mathematics Curriculum Standards and are renowned for their emphasis on contextual creation, inquiry-based learning, and cognitive adaptability. These textbooks represent the mainstream educational materials following China's "New Curriculum Reform." Textbooks from Prosveshcheniye Publishing House, Russia's largest and most authoritative publisher of basic education materials, are widely used nationwide. This edition inherits the long-standing traditions of Soviet-Russian mathematics education and is

renowned for its logical rigor, systematic theoretical framework, and depth of content. Selecting these two editions signifies the study of "benchmark" products within the mainstream educational systems of both countries. The comparative conclusions thus drawn possess high representativeness and universality, authentically reflecting the typical characteristics of mathematics education in China and Russia.

The seventh-grade upper-semester textbook published by Jiangsu Phoenix Science and Technology Publishing House is hereinafter referred to as the SJ edition textbook, and the sixth-grade lower-semester textbook published by Prosveshcheniye Publishing House is hereinafter referred to as the PR edition textbook.

Characteristics of the educational systems in China and Russia: China has a six-year primary school system, a three-year junior high school system, and a three-year senior high school system; Russia has a four-year primary school system, a five-year junior high school system, and a two-year senior high school system.

Table 1. The specific information selected for textbooks in both countries.

Country	Title	Main Author	Publisher	Publication Date
China	Compulsory education textbook mathematics grade 7 Part 1	Bao Jiansheng	Jiangsu Phoenix Science and Technology Press	2024
Russia	Mathematics Grade 6 Textbook Part 2	N. Ya. Vilenkin, V. I. Zhokhov et al.	Просвещение	2023

2.2. Selection of Research Content

Although the educational systems of the two countries differ slightly (China's seventh-grade first semester corresponds to Russia's sixth-grade second semester), the students are roughly the same age (approximately 13-14 years old) and are at a critical stage where the number system expands from arithmetic to algebra. The "introduction of rational numbers" represents a shared core and challenging component in both curricula. This alignment provides a solid foundation for content synchronization and cognitive development comparability in comparative research. Both countries' textbooks must address the fundamental questions: "How can students grasp the meaning of negative numbers?" and "How can a complete number system (the set of rational numbers) be constructed?" Examining their divergent approaches to this shared problem most effectively reveals underlying differences in instructional philosophy. Since mathematical definitions are crucial for conceptual understanding, this study focuses on the expansion of the number system—a logically interconnected topic in secondary mathematics—specifically analyzing the construction process of the rational number concept through comparative analysis of SJ and PR textbooks. The study will concentrate on comparing the presentation of core concept

definitions, including positive and negative numbers, integers and rational numbers, the number line, opposite numbers, and absolute values.

3. Comparison and Analysis

3.1. Comparison of the Knowledge System and Sequence of Rational Number Concepts

The SJ edition textbook places the content related to rational number concepts in Chapter 2 of the first semester of seventh grade, while the PR edition textbook places this content in Chapter 1 of the second semester of sixth grade (Chapter 4 of the first and second semester textbooks).

Table 2.The specific sequence of rational number concepts in the two editions of the textbook

Sequence	SJ Edition Textbook	PR Edition Textbook
1	2.1 Positive numbers, negative numbers, integers, and rational numbers	4.24 Positive numbers, negative numbers, and the number line
2	2.2 number line	4.25 Opposite numbers and integers
3	2.3.1 absolute value	4.26 mold
4	2.3.2 opposite number	4.35 rational number

Note: The "model" in the PR edition textbook is essentially the absolute value in the People's Education Press edition textbook. In the SJ edition textbook, the concepts of absolute value and opposite number are compiled in Section 2.3. For the sake of comparison and actual teaching arrangements, this article divides them into 2.3.1 and 2.3.2.

A comparative analysis of Table 2 reveals that the two editions of textbooks are highly consistent in terms of the scope of coverage of rational number concepts, but there are significant differences in the organizational structure and presentation order of the knowledge points. This study will focus on two dimensions for in-depth discussion: first, a systematic comparison of the similarities and differences in the definitions of relevant concepts in the two editions of textbooks; second, an in-depth analysis of the underlying causes of differences in the order of knowledge points and their pedagogical significance.

3.2. Comparative Analysis of Definitions of Related Concepts of Rational Numbers

Below, we will conduct a systematic comparative analysis in the order in which the relevant concepts are presented in the SJ edition textbook.

3.2.1. Positive and Negative Numbers

The two editions of the textbook are consistent in their mathematical essence of positive and negative numbers, both using symbols to distinguish the positive and negative properties of numbers, but there are differences in their expressions. The definition of positive and negative numbers in the SJ edition textbook has the following characteristics:

(1) Concise and clear expression, using concrete examples (such as noting that the symbol characteristic of positive numbers is no symbol or "+", and that of negative numbers is "−") to intuitively present the concept;

(2) Diverse example types, covering various numerical forms such as integers and decimals (e.g., +40,000, 1.7, 8,848.86), highlighting the broad applicability of the definition;

(3) International terminology is used, clearly providing the English terms "positive number" and "negative number," enhancing the textbook's international appeal.

In contrast, the PR version of the textbook's definition:

(1) Emphasizes formal characteristics, such as directly defining positive numbers as "numbers with a '+' sign";

(2) The examples are limited to integer forms (e.g., +3, +5) and do not include other types such as decimals;

(3) International terms are not labeled.

Table 3. Concept presentation

SJ Edition Textbook	PR Edition Textbook
Numbers such as 8848.86, 4, +40,000, and 1.7 are positive numbers; numbers such as − 80.97, − 6, −10,000, and−0.6 are negative numbers.	Numbers with the sign "+" are called positive and are written as +3, +5. Numbers with the sign"-"are called negative and are written as −3, −5.

The core commonality between the two versions of the textbook lies in their definitions of positive and negative numbers based on the external symbolic characteristics of numbers. The main differences lie in the richness of examples, the labeling of international terms, and the level of detail in the definitions. Overall, the SJ edition textbook maintains mathematical rigor while emphasizing teaching practicality and an international perspective, which helps students build a more comprehensive understanding of number concepts; the PR edition textbook, on the other hand, emphasizes the formal characteristics of concepts, reflecting a different editorial philosophy.

3.2.2. Integer

There are significant differences between the SJ edition and PR edition textbooks in terms of how they define the concept of integers: The SJ edition uses an extensional definition method, explicitly defining the elements of the integer set through enumeration (positive integers, zero, negative integers). This definition method is intuitive and clear, making it easier for beginners to quickly grasp the concept's extensional meaning, but it relatively weakens the intrinsic connections between mathematical concepts.

The PR edition textbook employs an intension-based definition method, constructing the concept by revealing the logical connection between integers and natural numbers. Specifically, it takes the set of natural numbers as its foundation (in the Russian number system, 0 is not part of

the natural numbers), uses opposite number operations to achieve algebraic expansion, and supplements the zero element to complete the construction of integers. This method emphasizes the generative and systematic nature of mathematical concepts but places higher demands on learners' abstract thinking abilities.

These two distinct writing paradigms reflect two typical orientations in mathematics education: the SJ edition emphasizes the acceptability of knowledge, while the PR edition emphasizes the logical rigor of mathematics.

Table 4. Integer Concept presentation

SJ Edition Textbook	PR Edition Textbook
Among positive numbers, numbers like $+7{,}998$ are called positive integers; among negative numbers, numbers like -9, -998 are called negative integers. Positive integers, negative integers, and zero are collectively referred to as integers. Integers and zero are what we commonly refer to as natural numbers.	Natural numbers, their opposite negative numbers, and zero are called whole numbers.

3.2.3. Rational Number

There are significant differences between the SJ edition textbook and the PR edition textbook in their definitions of rational numbers.

Table 5. Rational Number Concept presentation

SJ Edition Textbook	PR Edition Textbook
Integers and fractions are collectively referred to as rational numbers.	The number that can be expressed as $\frac{p}{q}$,p an integer, and q a natural number, is called a rational number.

The SJ edition defines rational numbers as "the collective term for integers and fractions." This definition is intuitive and helps students grasp the formal characteristics of rational numbers at an early stage. However, since "fractions" lack a clear definition, it is easy for students to become confused about the essence of rational numbers. For example, are $\frac{\pi}{2}$ and $\frac{\sqrt{3}}{2}$ considered "fractions"? If they are considered fractions, it would lead to the erroneous conclusion that they are rational numbers. More seriously, if we let x be any real number, since $x = \frac{x}{1}$, this definition would lead to the fallacy that "any real number is a rational number." This unrestricted definition of "fractions" has caused significant confusion for both teachers and students. The fundamental issue lies in the lack of restrictions on the number domains of the numerator and denominator. In contrast, the definition of rational numbers in the PR edition textbook (which restricts the numerator to integers and the denominator to natural numbers) offers the following advantages:

(1) Rigorous and unambiguous definition: The explicit specification of the range of values for the numerator and denominator allows clear determination based on the definition that $\frac{\pi}{2}$ and $\frac{\sqrt{3}}{2}$ are not rational numbers, effectively eliminating counterexamples.

(2) Revealing the logic of number field expansion:

The definition of rational numbers in the PR edition textbook lays the groundwork for a three-layer construction framework,

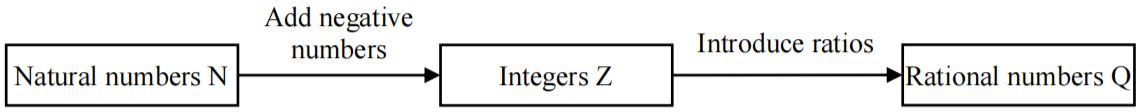

Figure 1. The logic of number field expansion

The system demonstrates the hierarchical construction process of mathematical concepts from simple to complex. This approach focuses on revealing the intrinsic logical chain of number field extensions, guiding students to understand the "construction process of mathematical objects," grasping the essential nature of the generation of mathematical concepts, and facilitating the establishment of core mathematical ideas regarding number field extensions.

(3) Laying the foundation for subsequent learning: This definition provides the conceptual foundation for constructing the rational number field (fractional field) from the integer ring in subsequent algebra and for fractional representations in number theory.

Comparative analysis shows that mathematical textbook definitions must balance rigor (eliminating counterexamples) with logical coherence (reflecting the evolution of concepts). The definition of rational numbers in the PR edition textbook is not only a model of mathematical rigor but also an excellent example of cultivating students' mathematical constructive thinking—it explains rational numbers within the broader framework of field extension. This exposition transcends formalistic descriptions, enabling students to see not isolated concepts but the logical trajectory of the evolution of the entire number system.

3.2.4. Number Axis

The definitions of the number line in the SJ edition and PR edition textbooks are identical in terms of mathematical content: both define it as a straight line with three elements—an origin, a positive direction, and a unit length.

Table 6. Number Axis Concept Presentation

SJ Edition Textbook	PR Edition Textbook
A straight line with a fixed origin, positive direction, and unit length is called a number line (number axis).	The line on which the origin, unit segment, and direction are selected is called the coordinate line. The number that shows the position of a point on the line is called the coordinate of that point.

Note: Unless otherwise specified in the following text, for convenience of expression, the coordinate axes in the PR edition textbooks are sometimes referred to as number lines.

The main differences between the two editions are as follows:

(1) Different sequencing: The SJ edition textbook places the number line after the study of rational numbers, teaching it as an independent section; The PR edition textbook integrates the number line with positive and negative numbers into the same section.

(2) Different functional focuses: The SJ edition textbook emphasizes "representing numbers with points," using points on the line to represent rational numbers. Its core objective is to establish a geometric representation of numbers, leveraging the number line's intuitiveness to aid understanding of rational number concepts, thereby embodying the principle of combining numbers and shapes. The PR edition textbook emphasizes "using numbers to represent points," stressing the use of numbers to indicate the position of points. Its focus is on constructing a coordinate system for points, placing greater emphasis on the geometric applications of rational numbers.

Overall, introducing the number line as a geometric representation and application tool for numbers aligns with students' cognitive patterns and facilitates an intuitive understanding of positive and negative numbers. Both editions utilize the number line to highlight the two dimensions of positive and negative numbers: the "essence" (understanding of the conceptual core) and the "application" (practical application skills). The distinctive features of both versions in terms of arrangement and functional positioning are worthy of mutual reference and dialectical unity, aiming to help students grasp the concepts of positive and negative numbers more comprehensively and deeply, enhance their application abilities, and thereby optimize the logical coherence and practical effectiveness of the textbook system.

3.2.5. Opposite Number

The two editions of the textbook are highly consistent in their definitions of the concept of opposite numbers. The main difference lies in the order of arrangement of the number line content: the SJ edition places it after the study of rational numbers, while the PR edition places it before the introduction of rational numbers. This difference in order reflects the underlying differences in their functional orientations.

Table 7. Opposite Number Concept Presentation

SJ Edition Textbook	PR Edition Textbook
Two numbers that differ only in sign are called opposites.	Two numbers that differ only in sign are called opposites.

SJ Edition Textbook (Rational Numbers → Opposite Numbers): This approach aims to use the concepts of opposite numbers and absolute value to highlight the core structure of rational numbers—the "sign" and "length" (absolute value)—enabling students to thoroughly understand

the composition of rational numbers. This path follows a logical progression from the general (integers) to the specific (the property of opposite numbers), emphasizing structural analysis of rational numbers.

PR Edition Textbook (opposite numbers → integers): In addition to the structural analysis function mentioned above, its core role lies in using the concept of opposite numbers to construct (or define) the set of integers. This highlights the foundational role of opposite numbers in the expansion of the number system, following a logical progression from the specific (opposite number relationships) to the general (constructing the set of integers), with a focus on the generation and expansion of the number system.

3.2.6. Absolute Value

Both the SJ edition and PR edition textbooks define the core concept of absolute value using a geometric interpretation: the distance between the point representing the number on the number line and the origin. This definition is intuitive and visual, making it easier for students to establish a foundational understanding when first encountering the concept.

Table 8. Absolute Value Concept presentation

SJ Edition Textbook	PR Edition Textbook
Generally speaking, the distance between a point representing a number on a number line and the origin is called the absolute value of that number. The absolute value of a number is denoted as |a| and read as "the absolute value of a."	The module of the number n is defined as the distance (in unit segments) from the origin to point N(n). It is written as $\|n\|$. The definition of the modulus can also be written as follows: $\|n\| = n, if\ n > 0,$ $\|n\| = 0, if\ n = 0,$ $\|n\| = -n, if\ n < 0,$

However, the PR edition textbook is more rigorous in its conceptual formulation. It explicitly specifies that this "distance" is the result of measuring "in units of length," rather than an actual physical measurement. This clarification effectively avoids ambiguity—for example, if the same distance is measured using centimeters and decimeters as units, the numerical representations will inevitably differ. This rigor reflects the PR edition textbook's approach to rational number instruction: guiding students to view rational numbers as independent, abstract algebraic objects rather than geometric entities. This approach is evident in its definition method: it not only provides a geometric explanation but also supplements it with the algebraic definition of absolute value.

In contrast, the SJ edition textbook adopts only the geometric meaning as the sole formal definition. The potential issue with this singular geometric definition approach is that it may blur the essential distinction between "numbers" (algebraic objects) and "distance" (metric quantities).

Although the two are highly compatible in the number line model, they belong to different conceptual categories. The essence of the integration of numbers and shapes lies in establishing connections rather than confusing categories. Anchoring the formal definition of algebraic concepts (numbers) entirely to geometric concepts (distance) may facilitate intuitive entry, but it may hinder students' deep understanding of absolute value as an independent, abstract algebraic object.

4. Conclusions and Implications

4.1. Conclusions

Through a comparative analysis of the rational number-related concepts in the SJ and PR editions of textbooks, it was found that the two editions are highly consistent in their coverage of core concepts, with the main differences lying in the way concepts are defined and the logical sequence in which they are presented. These differences fundamentally reflect two distinct textbook development philosophies: the PR edition emphasizes the logical consistency of mathematical knowledge and the intrinsic connections between concepts, focusing on the constructive logic of number system expansion; whereas the SJ edition prioritizes the intuitive presentation of knowledge and its immediate applicability, aiming to reduce cognitive load for beginners.

The comparative analysis indicates that the PR edition textbook has significant advantages in terms of the rigor of concept expression and the structural integrity of the knowledge system, providing important reference value for the compilation and instructional design of the SJ edition textbook. Meanwhile, the successful experiences of the SJ edition in terms of the intuitive nature of concept introduction and cognitive adaptability also offer valuable insights for the improvement of the PR edition textbook.

4.2. Implications

The definition of rational number-related concepts should not only focus on their representational forms but also reveal their underlying logical relationships. Based on the comparative analysis of the two versions of textbooks and their respective advantages, the following three suggestions are proposed for the presentation of rational numbers and their sequence concepts:

4.2.1. Optimizing the Presentation Sequence of Related Concepts

Suggested sequence: positive numbers, negative numbers, 0 → number line → opposite numbers → integers → rational numbers → absolute value. Placing the number line after positive and negative numbers and before rational numbers facilitates students' geometric understanding of positive and negative numbers. Introducing opposite numbers before integers helps reveal the constructive logic of expanding the number system.

4.2.2. Clarify the functional positioning of concepts

When introducing concepts, emphasize their core functions:

(1) Number line: serves as the geometric representation of rational numbers, achieving the integration of numbers and shapes, and aiding in the understanding of positive and negative numbers.

(2) Opposite numbers: highlight the symbolic characteristics of rational numbers, serving as a key link in the expansion of the natural number system to the integer system.

(3) Integers: form the basic elements of the rational number system.

4.2.3. Strengthening the Rigorous Definition of Concepts

(1) Definitions of positive and negative numbers: It is recommended to use property-based definitions (e.g., positive numbers are numbers greater than 0).

(2) Integers: The collective term for natural numbers and their opposite numbers.

(3) Rational numbers: Numbers that can be expressed in the form of a fraction p/q (where p is an integer and q is a non-zero integer).

(4) Absolute value: It is recommended to use an algebraic definition as the main approach, with a geometric interpretation as a supplement.

Through a comparative analysis of the rational number concepts in the SJ and PR editions of textbooks, this study clearly highlights the differences in the two textbooks' editorial philosophies and their respective values: the PR edition excels in logical rigor and system construction, while the SJ edition excels in intuitive introduction and cognitive adaptation. This difference is not a matter of superiority or inferiority but provides complementary perspectives for textbook optimization. Based on this, the study further proposes specific recommendations for optimizing the sequence of concept presentation, clarifying core functional positioning, and strengthening the rigor of definitions. The core of these recommendations lies in guiding textbook compilation and teaching practices to place greater emphasis on revealing the intrinsic logical connections of concepts while also considering students' cognitive starting points. Future rational number concept instruction should actively draw on the essence of both editions of the textbook, seeking a better balance between logical depth and cognitive breadth to promote students' formation of a profound and accurate understanding of the rational number system.

Author Contributions:

Conceptualization, H.W.; methodology, H.W; validation, H.W; formal analysis, H.W; writing—original draft preparation, H.W; writing—review and editing, H.W; visualization, H.W; supervision, H.W; project administration, H.W; funding acquisition, H.W. All authors have read and agreed to the published version of the manuscript.

Funding:

This research was funded by The 15th Issue of Jiangsu Province Primary and Secondary School Teaching Research Project（Grant number:2023JY15-GX-L25).

Institutional Review Board Statement:

Not applicable.

Informed Consent Statement:

Not applicable.

Data Availability Statement:

Not applicable.

Acknowledgments:

Not applicable.

Conflict of Interest:

The authors declare no conflict of interest.

References

Alexander, K., & Vogeli, B. (2010). Russian mathematics education: history and world significance. World Scientific Publishing.

Cai, J., & Jiang, C. (2017). An Analysis of Problem-Posing Tasks in Chinese and US Elementary Mathematics Textbooks. International Journal of Science and Mathematics Education, 15(8), 1521-1540.

Fan, L., Zhu, Y., & Miao, Z. (2013). Textbook research in mathematics education: development status and directions. ZDM, 45(5), 633-646.

Karp, A. (2007). Russian mathematics education: History and world significance. International Journal for the History of Mathematics Education, 2(1), 43-60.

Li, Y. (2004). A comparison of mathematics curricula in China and the United States. Zentralblatt für Didaktik der Mathematik, 36(1), 1-13.

Sfard, A. (2007). When the Rules of Discourse Change, but Nobody Tells You: Making Sense of Mathematics Learning From a Commognitive Standpoint. Journal of the Learning Sciences, 16(4), 565-613.

Shimizu, Y. (2009). Japanese approach to teaching mathematics via problem solving. In B. Kaur (Ed.), Mathematical Problem Solving: Yearbook 2009, Association of Mathematics Educators (pp. 89-101). World Scientific.

Van Hoof, J., Verschaffel, L., & Van Dooren, W. (2017). Number sense in the transition from natural to rational numbers. British Journal of Educational Psychology, 87(1), 43-56.

Vygotskij, L. S. (1979). Mind in society: The development of higher psychological processes. Harvard University Press.

Wu, H. (2011). The mis-education of mathematics teachers. Notices of the AMS, 58(3), 372-384.

iEducation, 2025, 1(2), 61-73
https://doi.org/10.71204/eytj1362

A Study on Implementation Pathways for AI-Powered Full-Process Intelligent Courses to Enhance Precision in Medical Education

Xiaozhong Chen [1], Jianing Liang[2], Yueyang Jiang [3], Xiaofeng Jin [4, *]

[1] School of Graduate, Wenzhou Medical University, Wenzhou 325000, China

[2] School of Traditional Chinese Medicine, Wenzhou Medical University, Wenzhou 325000, China

[3] School of medical humanities and hospital management, Wenzhou Medical University, Wenzhou 325000, China

[4] Teade Union, Wenzhou Medical University, Wenzhou 325000, China

*** Correspondence:**

Xiaofeng Jin

wmcjxf@163.com

Received: 29 September 2025 /Accepted: 7 November 2025 / Published online:14 November 2025

Abstract

This study focuses on innovative pathways for leveraging artificial intelligence (AI) technology to enhance higher education course development, with the core objective of establishing a smart course system that encompasses the entire process from course design, teaching implementation, learning support, to evaluation and feedback. Leveraging AI technologies such as knowledge graphs and large language models, a systematic smart course platform has been developed, featuring core modules including AI-assisted lesson plan design, intelligent student performance analysis, digital virtual teachers, adaptive learning pathways, and intelligent assessment feedback. Practical applications have demonstrated that the platform addresses key challenges through four pillars: 'AI-empowered course design and planning,' 'innovative "teacher/student/machine" deep interaction teaching models,' 'adaptive learning paradigms enabling "active learning,"' and 'enhancing assessment methods through diversified evaluation.' This has effectively addressed key issues such as inaccurate learning situation analysis, monotonous teaching models, insufficient teacher AI skills, and the lack of dynamic quantification in evaluations. It has significantly improved course teaching quality and student learning efficiency, providing a scalable smart course model for the digital transformation of higher education.

Keywords: Artificial Intelligence; Smart Courses; Full-Process Construction; Knowledge Graphs; Adaptive Learning; Medical Education

1. Introduction

In the era of artificial intelligence driving the digital innovation and development of education, higher education is undergoing profound transformation. In 2024, the Higher Education Division of the Ministry of Education issued a document titled *Artificial Intelligence Driving the Digital Innovation and Development of Higher Education*, clearly stating the need to promote the deep integration of artificial intelligence and other information technologies with education and teaching, and to implement an artificial intelligence empowerment initiative (Wang et al., 2024). For medical schools facing the challenge of balancing the expansion of graduate education scale with the enhancement of educational quality, leveraging AI technology to strengthen the construction of online medical and pharmaceutical courses and promote educational reform has become an urgent need to improve educational quality. Current course development faces systemic challenges that hinder improvements in teaching quality, primarily manifested in four key issues: inaccurate analysis of student learning situations, insufficient student initiative, monotonous teaching models, and outdated evaluation systems (Thurzo, 2025). This paper will address these issues by adopting a problem-oriented approach to establish an effective, multi-dimensional AI-based course quality improvement system and explore a comprehensive smart course development implementation pathway tailored to medical education. This holds significant implications for improving the teaching quality of medical courses at medical institutions. However, the model and implementation path require a larger dataset and broader application scope during the initial phase to assess their rationality and effectiveness. This will also provide the most direct information and data sources for subsequent model optimisation and path refinement, thereby enhancing the efficacy of the model and path. This holds significant implications for improving teaching quality, ensuring students' comprehension of course content, and fostering their innovative thinking.

2. Current Status and Challenges of AI Integration in Medical Education

The AI revolution is sweeping the globe, and medical education—a field characterised by its knowledge-intensive nature and stringent skill requirements—is on the brink of a silent transformation. AI-enabled medical education is hailed as the ‘golden key’ to breaking free from traditional constraints and reshaping the future model of physician training. Currently, the integration of AI and medical education exhibits a pattern of ‘localised application enthusiasm but systemic integration lag.’ According to data from Grand View Research, AI-driven virtual anatomy platforms, intelligent diagnostic simulation systems, and medical imaging assistance tools have already been implemented in some top medical schools. The global medical education AI market is projected to reach 6.18 billion US dollars by 2027. However, its technological applications remain fragmented, such as surgical simulation training and knowledge graph-assisted learning, and have not yet formed an ecosystem-level integration covering the entire chain of ‘teaching, learning, assessment, and management.’ Meanwhile, educational administrators are beginning to recognise the value of learning behaviour data, but data silos remain widespread, and utilisation efficiency remains low. A 2023 survey by China Medical Education Technology covering 50 medical schools showed that only about 15% of institutions

have established preliminary unified learning data analysis platforms (Huang et al., 2023). This indicates that the current integration process is still in its exploratory phase, facing numerous challenges, yet also holding significant untapped potential.

Through daily teaching experiences and surveys of key groups within campuses, combined with existing data and literature, the team identified four major issues in the current integration of AI technology with medical education: First, there is the issue of inaccurate learning situation analysis. AI could integrate multi-dimensional data such as online learning duration, resource click streams, forum interactions, simulation operation paths, and even eye movements and physiological signals (in the experimental stage). However, in reality, data sources are fragmented, and algorithm models are simple, making it impossible to construct accurate personalised learning profiles, resulting in teaching interventions that are 'aimless.' Currently, medical schools that have introduced AI learning platforms primarily use them for automatically tracking video viewing duration, failing to delve into the correlation between behavioural patterns and knowledge mastery. The second challenge is insufficient learning initiative, with existing issues such as passive information dissemination and lack of personalisation remaining severe (Zhou et al., 2025). Current AI tools often feature rigid designs, awkward interactions, and a lack of deep appeal and emotional connection, resulting in a poor user experience for students and an inability to effectively ignite learning enthusiasm. The third challenge is the extreme lack of diversity in current teaching models. Innovative teaching models, whether due to technical costs, insufficient digital literacy among teachers, or outdated teaching designs, have failed to effectively integrate into mainstream classrooms, leaving the traditional 'lecture + memorisation' model largely intact (Fu, 2025). Finally, the lagging evaluation system continues to have a significant impact, with evaluation reforms facing multiple challenges such as technical reliability verification, ethical and privacy concerns, and difficulties in integrating with traditional scoring systems, resulting in slow progress in implementation. The severe lag in evaluation feedback makes it difficult for students to receive timely and effective guidance for improvement. This issue has seriously impacted the integration of AI technology and medical education (Lang et al., 2025).

To ensure that the aforementioned AI-enabled smart course platform (covering its technical architecture based on knowledge graphs and four core functional systems) can be effectively implemented and truly solve the deep-seated issues in course development, this study has constructed a clear and collaborative implementation model and working mechanism. This model focuses on addressing the key bottlenecks in current teaching practices, driving systematic improvements in course quality through the deep application of the platform's core capabilities and multi-stakeholder collaboration.

3. Establishing a closed-loop quality improvement mechanism for AI-enabled medical and pharmaceutical courses

To ensure that the aforementioned AI-enabled smart course platform (covering its technical architecture based on knowledge graphs and four core functional systems) can be effectively implemented and truly solve the deep-rooted problems in medical course development, this study

has constructed a clear and coordinated implementation model and working mechanism. This model focuses on breaking through the key bottlenecks in current medical teaching practices, driving systematic improvements in medical course quality through the in-depth application of the platform's core capabilities and multi-party collaboration.

3.1. Construction of the Smart Course Platform Implementation Model and Analysis of Its Mediating Role

3.1.1. Establishing New Pathways for Medical Education Logic and Competency Development with a Multidimensional Dynamic Knowledge Graph as the Core Foundation

Current medical education practices commonly face two critical challenges: firstly, insufficient precision in learner profiling hinders systematic understanding of individual and group cognitive characteristics and competency disparities among medical students, thereby undermining evidence-based differentiated instruction (Jin, 2025); Secondly, course resource development often lacks systematic planning, resulting in fragmented resource types that poorly align with teaching objectives, thereby failing to effectively support the cultivation of advanced medical competencies (Fu et al., 2025). To address these challenges, this platform utilises a proprietary multi-dimensional dynamic knowledge graph as its core foundation. By deeply integrating the unique educational logic and competency development pathways of medical disciplines, it constructs a comprehensive precision teaching support system.

This knowledge map is not a static structure in the traditional sense, but rather integrates three core dimensions: firstly, the advanced problem-solving framework unique to medical disciplines, encompassing complex pathways from fundamental mechanisms to clinical applications; secondly, a competency-based dual-track map of clinical and research capabilities, clarifying the operational skills, cognitive approaches, and research literacy required at different learning stages; Thirdly, a continuous thread of medical humanities and professional ethics course-based ideological and political education, embedding elements such as ethical decision-making, doctor-patient communication, and professional ethics within the knowledge structure. This dynamically updated, multi-dimensional knowledge graph system provides a theoretical model and data annotation framework for conducting in-depth learning analytics.

This platform's multidimensional dynamic Knowledge Graph is constructed based on the unique logic of Medical education: it first integrates multi-source data including the Chinese Undergraduate Medical Education Standards, authoritative textbooks, clinical cases, and teaching processes. In its construction methodology, it first distils the discipline-specific problem frameworks, competency objectives, and core medical humanities themes. Subsequently, machine extraction forms a structured Knowledge Graph with precise associations between ‘knowledge nodes – competency tiers – ideological and political elements’. Leveraging data feedback and periodic revisions for dynamic iteration, it both responds to updates in medical knowledge and adapts to the competency development needs of medical students at different stages, providing a dynamically evolving knowledge foundation for precision teaching.

To address the significant challenge posed by insufficient precision in learning analytics, which hinders systematic understanding of individual and group cognitive characteristics and ability

disparities among medical students, the task of constructing a multidimensional dynamic Knowledge Graph is now imperative.Leveraging this knowledge map, the platform has developed an intelligent learning diagnostics tool for educators. This tool integrates behavioural data from multiple learning contexts—theoretical study, simulation training, and clinical placements—to generate individual and cohort-level learning reports through node-based mapping and competency alignment. Educators can not only pinpoint students' weaknesses at specific knowledge nodes—such as insufficient spatial comprehension of anatomical structures or delayed reasoning abilities regarding certain pathological mechanisms—but also identify collective biases in clinical reasoning, research literacy, or doctor-patient communication. This enables truly precise identification, providing clear direction for subsequent instructional design.

In addressing the challenges of course resource development—which often lacks systematic planning, features fragmented resource types, and exhibits poor alignment with teaching objectives，the technical team collaborates with medical educators to comprehensively review and optimise existing teaching resources under the systematic guidance of the knowledge graph. This effort also involves planning and constructing a new generation of multimodal medical teaching resource repositories. Resource development strictly adheres to an integrated ‘objective-resource-assessment’ principle, striving for high alignment with medical education goals and student cognitive characteristics. Specifically, the platform prioritises developing high-quality, diverse teaching resources such as: - Interactive 3D models precisely matching complex anatomical structures, enabling multi-angle observation and virtual dissection; - Virtual simulation experiments illustrating dynamic pathological mechanisms, aiding comprehension of disease onset and progression; - Case libraries integrating authentic clinical scenarios, covering end-to-end training from diagnostic reasoning to treatment planning. Additionally, it incorporates evidence-based research training modules and standardised patient simulation resources.

Each resource category is precisely linked to corresponding knowledge nodes, competency objectives, and ideological elements within the knowledge graph. For instance, a 3D model of a cardiac valve lesion connects not only to anatomical knowledge nodes but also aligns with clinical differential diagnosis competency training objectives, while integrating the ideological discussion point of ‘patient informed consent’ from medical ethics. This structured, tag-based resource management enables the platform to automatically recommend personalised resource combinations tailored to students' varying competency levels and learning styles based on learning analytics. This supports tiered teaching and individualised guidance by educators.

Ultimately, through a knowledge graph-driven closed-loop system of learning analytics and resource development, the platform significantly enhances the targeted nature of medical education resource construction and the precision of instructional design. Educators can formulate tailored teaching strategies based on objective data, while students gain learning experiences better aligned with their individual cognitive development pathways. This fundamentally propels medical education from a ‘one-size-fits-all’ approach towards ‘personalised cultivation,’ thereby supporting the systematic development of high-calibre medical professionals.

3.1.2. Establish an intelligent platform to serve as an intermediary in addressing the teaching and learning needs of both staff and students.

To overcome challenges such as the limited diversity of teaching methods in medical curricula (particularly in foundational and bridging disciplines like anatomy, pharmacology, and pathology), restricted depth and breadth of interaction between medical students and faculty, and difficulties in sharing high-quality pedagogical experiences (Yang et al., 2025), the intelligent teaching tools provided by the platform serve as a critical support. Medical educators can flexibly utilise features such as AI lesson plan design, automated courseware generation and enhancement, and AI content creation (e.g., generating teaching scenario descriptions based on real cases, expanding clinical decision analysis) to efficiently construct high-quality, highly interactive teaching materials. Concurrently, digital virtual tutors (AI teaching assistants) provide round-the-clock intelligent medical knowledge Q&A and guidance. AI-accelerated video production (such as rapid creation of surgical procedure breakdown videos and disease progression animations) significantly lowers the threshold for high-quality video resources. Combined with intelligent interactive classrooms (incorporating technologies like universal whiteboards), this creates a new ecosystem for medical digital-intelligent classrooms featuring seamless online-offline integration and deep ‘teacher-student-machine’ collaboration. This process is complemented by the establishment of virtual teaching and research rooms, effectively facilitating the consolidation and sharing of high-quality teaching experiences, resources, and strategies among medical educators through intelligent tools. This includes effective methods for teaching complex diseases and clinical reasoning training, greatly enriching teaching formats, enhancing classroom engagement and efficiency, and effectively addressing the challenges of teaching abstract medical concepts while overcoming limitations in class hours and student comprehension capacity.

However, as the core implementers, medical educators' proficiency in intelligent teaching skills and application levels directly impacts the platform's effectiveness. To address this, the platform itself incorporates low-threshold AI-assisted tools (such as intelligent generation of medical examination questions, automated analysis of clinical skills simulation operation data, and automatic generation of learning progress reports), providing an intuitive ‘scaffolding’ for teachers to develop their digital literacy. Concurrently, the institution provides systematic training and technical support at the organisational level. This focuses on empowering educators to efficiently evaluate AI-recommended medical resources, interpret learning data and assessment profiles to inform teaching decisions, and master AI-integrated instructional design principles and methodologies. This significantly lowers the threshold for AI adoption, enabling educators to concentrate on core medical teaching design and student guidance, thereby facilitating a seamless transition from traditional to intelligent teaching.

3.2. Data-Driven Dynamic Evaluation and Faculty Collaborative Development Mechanism

Traditional medical education evaluation systems have long faced multiple challenges, including lagging assessment timelines, limited assessment modalities, and subjective evaluation practices (Lillehaug & Lajoie, 1998). Conventional methods such as written examinations and practical assessments often lack real-time feedback and multidimensional quantitative support. Consequently, they struggle to comprehensively reflect students' knowledge acquisition, clinical

reasoning abilities, and professional competence levels (Li et al., 2024). Moreover, they fail to provide sustained, reliable data support for the dynamic optimisation of teaching processes (Chen et al., 2024). To fundamentally transform this situation, this platform has established a data-centric, intelligence-driven dynamic evaluation and collaborative development mechanism, achieving a systemic shift in medical education assessment from ‘experience-led’ to ‘data-driven’.

The core of this mechanism lies in establishing a closed-loop system covering the entire ‘teaching, learning, assessment, and optimisation’ process. Through embedded data collection modules, the platform comprehensively tracks and records medical students' behavioural data across multiple learning scenarios. This encompasses operational steps in virtual simulation experiments, clinical decision pathways in case simulations, the quality of contributions in online interactive discussions, and performance in various theoretical and practical assessments. These high-frequency, multi-modal, continuous data points form the digital footprint of student learning, providing a robust foundation for in-depth analysis.

At the intelligent assessment level, the platform integrates multiple cutting-edge artificial intelligence technologies to construct a robust medical education-specific evaluation system. For instance, the medical English oral assessment module not only evaluates pronunciation accuracy but also analyses the appropriateness of terminology usage and communication fluency. The automated marking of case analysis reports and clinical decision-making essays is achieved through an intelligent assessment system customised for Medical education scenarios. The system constructs its assessment logic by learning from two core corpora: firstly, representative error samples from students' routine assignments (e.g., common issues such as knowledge gaps); secondly, the latest clinical practice guidelines and textbooks (including standardised textbooks like Internal Medicine and Diagnostics, alongside industry standards such as the Chinese Medical Association's Guidelines for the Diagnosis and Treatment of Acute Myocardial Infarction and the Chinese Clinical Practice Guidelines for Community-Acquired Pneumonia in Emergency Settings), which define the boundaries of evidence-based medical knowledge and the basis for clinical decision-making. Furthermore, the platform's plagiarism detection and originality assessment capabilities for medical literature reviews significantly enhance the practicality of academic integrity education. Collectively, these capabilities expand the depth and efficiency of assessment, enabling comprehensive evaluation of ‘knowledge-skills-competencies’.

Leveraging multi-source data and intelligent assessment results, the platform's advanced learning analytics engine generates highly interpretable quantitative evaluation profiles and dynamic learning progress reports. These evaluation profiles encompass not only the internalisation of medical knowledge but also comprehensive competencies such as clinical reasoning ability, practical operational skills, doctor-patient communication performance, and medical ethical judgement. Each report visually presents individual and group learning progress, capability gaps, and developmental trends, enabling educators to move beyond reliance on subjective teaching judgements. Instead, they can make precise instructional decisions grounded in objective, comprehensive, and timely data insights.

Building upon this foundation, the platform effectively fosters teachers' professional development and optimises instructional practices. Educators can dynamically adjust teaching strategies based on real-time feedback from the system. For instance, they may provide focused explanations on common knowledge gaps within a cohort or design differentiated training content and resource delivery strategies tailored to students' varying ability levels. Concurrently, the platform supports data-driven teaching reflection and peer collaboration among educators, facilitating a continuous cycle of improvement: assessment → diagnosis → intervention → re-evaluation.

Concurrently, this mechanism enables the platform's self-iteration. Drawing upon continuously accumulated assessment feedback and learning efficacy data, the platform continually refines its internal recommendation logic and resource matching mechanisms. This includes adjusting learning pathway difficulty gradients, increasing the proportion of exercises targeting weaker areas, and introducing case studies and simulation training better aligned with students' current proficiency levels. Consequently, teaching resources are more precisely matched to individual needs.

Ultimately, this data-driven dynamic assessment mechanism, coupled with the collaborative development model for teaching staff, forms a virtuous 'assessment-feedback-optimisation' closed loop. This not only significantly enhances the timeliness and accuracy of medical education evaluation but also greatly strengthens the responsiveness and adaptability of the teaching system. It signifies our progression towards a new paradigm of medical education that is more inclusive, scientific, and growth-oriented, providing robust support for cultivating comprehensively skilled medical professionals.

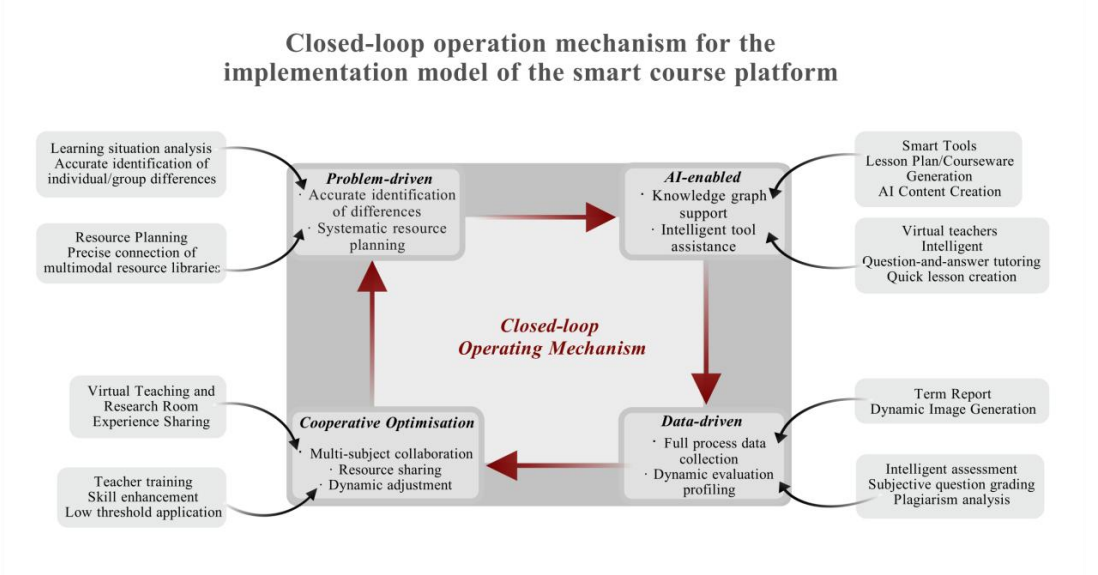

Figure 1. Diagram illustrating the closed-loop operation mechanism of the smart course platform implementation model

4. Research and Validation Methodology

To ensure the scientific rigour, efficacy, and educational applicability of the intelligent models within the constructed Smart Courses platform, this study employs a four-stage closed-loop

research methodology: ‘Model Construction – Data Collection – Effectiveness Validation – Iterative Optimisation’. This approach systematically advances the development of intelligent models and their educational validation.

4.1. Model Construction and Data Foundations

Centred upon a Medical education Knowledge Graph, the platform integrates natural language processing, machine learning, and adaptive learning algorithms to construct four core intelligent modules: learning diagnostics, resource recommendation, virtual tutoring, and intelligent assessment. The Knowledge Graph is built from multi-source data including the Chinese Undergraduate Medical Education Standards, authoritative medical textbooks, authentic clinical cases, and teaching logs. Through entity extraction, relationship mapping, and semantic fusion, it forms a three-dimensional dynamic structure encompassing ‘knowledge-competency-proficiency’. Model training utilises corpus and behavioural data specific to medical education scenarios, including student simulation pathways, case analysis texts, interactive discussions, and assessment records, ensuring the model's adaptability within the medical education domain. During the refinement phase, the platform underwent preliminary trials across medical colleges and medical schools within comprehensive universities within the province, yielding foundational data for subsequent model enhancement and iteration.

4.2. Design of Intelligent Model Validation Methods

To validate the model's efficacy, a combined internal and educational empirical verification mechanism was devised:

(1) Internal technical validation: Employing cross-validation and A/B testing to assess the model's performance in recognition accuracy, recommendation relevance, and feedback timeliness. For instance, the rationality of adaptive recommendation algorithms was tested by comparing the consistency between AI-generated learning pathways and expert-designed pathways in achieving teaching objectives.

(2) Educational Empirical Validation: Pilot classes were selected within partner medical colleges to conduct a semester-long comparative teaching experiment. The experimental group employed the Smart Courses platform for full-process teaching, while the control group adhered to traditional teaching methods. Comprehensive evaluation of the platform's practical effects—including enhancing student learning outcomes, strengthening teaching interaction, and optimising instructional decision-making—was achieved through multidimensional data analysis. This encompassed pre- and post-test score comparisons, learning behaviour analysis, teacher-student interviews, and satisfaction questionnaires.

4.3. Data Collection and Analysis

The platform incorporates an integrated data collection module that records real-time behavioural data throughout the entire learning process, encompassing resource clickstreams, task completion durations, interaction frequencies, assessment outcomes, and teaching adjustment records. Primary data sources are the participating teaching institutions and medical education offices, ensuring authenticity. Following data collection, the model classifies and analyses the

information using descriptive statistics, cluster analysis, and association rule mining to identify learner group characteristics, teaching bottlenecks, and optimisation directions. Concurrently, qualitative analysis methods are applied to code and analyse teacher interview content, extracting application patterns and barriers for intelligent tools in actual teaching practice to support system development.

Regarding data sources, foundational data primarily comprises 3,500 records from Wenzhou Medical University's relevant disciplines—including basic medicine, clinical medicine, nursing, pharmacy, and biomedical engineering—serving as the model's core dataset. This data predominantly originates from undergraduate medical and pharmaceutical students. The team employed stratified sampling based on year-level proportions to ascertain each discipline's willingness to utilise the online intelligent platform and identify fundamental requirements. Concurrently, feedback was gathered from subject lecturers across university departments and affiliated hospitals, yielding 580 additional data points. This combined input enabled the development of an intelligent course platform better aligned with both teaching requirements and student needs.

4.4. Model Optimisation and Scalability Assessment

Building upon issues and feedback identified during validation, the research team undertook iterative refinements to the model. Key enhancements focused on updating the Knowledge Graph mechanism, refining the granularity of learner profiling, and improving the resource tagging system. Concurrently, multiple pilot trials across diverse medical institutions tested the model's adaptability and scalability, evaluating its potential for deployment across varied teaching environments, student cohorts, and course types.

5. Conclusion and Outlook

The team identified four major issues encountered in daily teaching and dedicated itself to exploring and implementing AI-enabled solutions for the entire course development process. The core of this effort was the construction of an intelligent course platform based on a knowledge graph and integrating multiple AI technologies. Through the synergistic effects of four mechanisms—precision planning and resource development, intelligent teaching implementation and interaction, support for enhancing teachers' AI literacy, and data-driven dynamic evaluation and optimisation—the team effectively addressed the key bottlenecks in medical course development. The implementation framework has been largely established, but the model remains in a trial and refinement phase.The replicability of the system requires urgent enhancement. Different institutions exhibit varying degrees of granularity in their teaching platforms, with somewhat ambiguous functional positioning requirements. Consequently, a single system cannot adequately meet the needs of all medical schools. Even with the replication of basic procedures and algorithms, modifications remain necessary to accommodate the specific demands of different institutions. The precision of individual student coverage requires improvement. This stems from undergraduate education's holistic approach to student development, where platforms primarily target entire cohorts rather than individuals. While personalised sections exist, they

often function more as self-exploration tools, diminishing the effectiveness and utility of personalised modules. The evaluation metrics within the overall assessment framework remain insufficiently comprehensive and specific. Differences in assessment criteria across regions, identified during the pilot phase, necessitate ongoing refinement of the evaluation indicators. These should be categorised to complete the construction of the assessment system. Future research will address these existing issues by deepening efforts in the following directions: first, further expanding the application of AI in precise assessment of English proficiency (especially speaking and writing) and intelligent cultivation of cross-cultural communication skills; second, deepening algorithm optimisation for personalised learning paths to enhance the precision and adaptability of recommendations; third, expanding the research sample size and extending the research period to track the long-term impact of smart courses on students' learning outcomes and skill development; Fourth, explore pathways for integrating AI ethics education into smart courses. The team hopes that the smart course paradigm established in this study will not only provide a practical model for medical education but also serve as a reference for the digital transformation of higher education more broadly, contributing to the high-quality development of medical higher education. Additionally, by pioneering research in the current untapped area of constructing a fully smart course development platform using AI technology, the team aims to lay the groundwork and accumulate experience to provide more samples and insights for future studies.

Author Contributions:

Conceptualization, Xiaozhong Chen; methodology, Yueyang Jiang; software, Xiaozhong Chen; validation, Jianing Liang; formal analysis, Jianing Liang; investigation, Xiaozhong Chen; resources, Xiaozhong Chen; data curation, Yangyang Zhang.; writing—original draft preparation, Jianing Liang; writing—review and editing, Xiaozhong Chen; visualization, Xiaozhong Chen; supervision, Xiaofeng Jin; project administration,Xiaofeng Jin; funding acquisition, Xiaozhong Chen. All authors have read and agreed to the published version of the manuscript.

Funding:

This research was funded by ①National Research Project on the Construction and Teaching Research of Online Courses for Postgraduate Students in Medicine and Pharmacy: Research on the Development and Implementation Pathways of AI-Based Smart Courses Throughout the Entire Process, grant number:B_YXC2024-02-03_10.②The second batch of provincial teaching reform projects for graduate students in Zhejiang Province during the 14th Five-Year Plan：Research on the teaching reform of "curriculum ideology and politics" integrated into the training of medical graduate students from the perspective of healthy China，grant number:JGCG2024303.③National Research Project on the Construction and Teaching Research of Online Courses for Postgraduate Students in Medicine and Pharmacy: "Medicine" Heart Directed Towards the Party, People First -- Exploration and Practice of "Curriculum Ideology and Politics" Construction in Medical Schools,grant number:B_YXC2024-02-02_10.④General

Projects of the Zhejiang Provincial Department of Education:The Design and Exploration of the Curriculum Ideological and Political System in New Era Medical Colleges and Universities,grant number:Y202352876.⑤Construction project for online courses in Chinese medical studies for graduate students:Medical Humanities Course Group for Graduate Students in Pharmaceutical Sciences,grant number:A_YXC2024-01-01_10.

Institutional Review Board Statement:

Not applicable.

Informed Consent Statement:

Not applicable.

Data Availability Statement:

Not applicable.

Acknowledgments:

We incerely thank the journal for the opportunity to publish, and we recognize that the efforts of our research team and all the teachers and students involved have contributed to this achievement. It is an honor for our research team to be able to offer relevant insights and perspectives based on our experiences and the foundation of our institution in this field. In the future, we will continue to delve deeper into this area, hoping to make even greater contributions!

Conflict of Interest:

The authors declare no conflict of interest.

References

Chen, B. W., Wu, K. H., Wei, J., Wang, Y. Q., Li, Z. S. N., Yang, S., & Meng, X. Y. (2024). Practice and Thinking of the Supervisor Responsibility System in the Connected Training Mode of Undergraduates and Postgraduates of Medical Students. International Journal of Geriatrics, 45(01), 121-125.

Fu, H. R., Liu, W. J., Zhu, L. Y., Long, Y., Xiang, B., & Yang, H. J. (2025). Exploration of AI-Assisted Diagnosis Teaching Pathways Based on Huazhi Yihui Platform. Continuing Medical Education, (06), 112-115.

Fu, M. (2025). Exploration of a Knowledge Graph-based Smart Teaching Model for Transportation Management Courses. Auto Time, (17), 90-92.

Huang, Y. P., Wu, H. D., Liu, H., Dai, H. L., Jin, L., Wang, L. J., & Sun, J. (2023). Construction and practice of formative evaluation system from the perspective of educational informationization 2.0. China Medical Education Technology, (04), 458-462+467.

Jin, G. P. (2025). Exploratory practice of integrating medical humanities education into clinical teaching of specialised traditional Chinese medicine disciplines. Journal of Chinese Medicine Management, (14), 255–257.

Lang, L. L., Su, J. T., Ren, W. N., Meng, J. J., Lu, Z. Z., Shen, T., & He, H. M. (2025). Generative

AI Empowering Higher Medical Education: The Path of Transformation from Curriculum to Assessment. Chinese Journal of Medical Education Technology, 1-6.

Li, J., Li, Y. K., Wang, Q. B., Liang, Y. B., Luo, W. L., Chen, X. M., & Ke, Y. (2024). Preliminary Exploration of Education Effect of Innovation and Entrepreneurship Competition for Medical Students with the Theme of Medical-Engineering Integration. Medical Education Research and Practice, (04), 404-408.

Lillehaug, S. I., & Lajoie, S. P. (1998). Ai in medical education—another grand challenge for medical informatics. Artificial Intelligence in Medicine, (3), 197.

Thurzo, A. (2025). How is AI Transforming Medical Research, Education and Practice?. Bratislava Medical Journal, 126, 243–248.

Wang, F., Liu, Y. Q., & Zhou, T. H. (2024). Artificial Intelligence Driving Digital Innovation in Higher Education. Chinese Higher Education, (Z1), 9-12.

Yang, X. M., Shen, Z. F., Xiong, Y. B., Chen, W. D., Chen, H., Chen, Y. X., & Shao, Y. J. (2025). Research on artificial intelligence enabling medical students to be creative. Journal of Bengbu Medical University, (01), 46-51.

Zhou, Z., Cao, J., Zhou, B. B., Din, L., Zhu, H. Y., Guo, H. J., & Yang, B. (2025). Exploration and Practice of AI-empowered Pharmaceutical Education Reform. Pharmaceutical Education, (04), 1-6.

iEducation, 2025, 1(2), 74-95
https://doi.org/10.71204/158qy339

Flagships and Community Builders: A Configurational Analysis of High-Impact University Ideological and Political Education WeChat Accounts

Shuang Wang [1, *]

[1]School of Convention-Exhibition Economy and Management, Shanghai Institute of Tourism, Shanghai 201418, China

[*]**Corresponding Author**
Shuang Wang

wshroc@sitsh.edu.cn

Received: 15 October 2025 / Accepted: 11 Novembe2025 / Published online: 14 November 2025

Abstract

Organizational success on digital platforms rarely emerges from single "best practice" factors but from the complex interplay among strategic elements. Yet, research has predominantly relied on linear models ill-equipped to capture this causal complexity. Addressing this gap, this study draws on configuration theory and employs fuzzy-set Qualitative Comparative Analysis (fsQCA) to uncover the pathways to high impact for 30 top-performing university Ideological and Political Education (IPE) WeChat accounts. Our analysis reveals two crucial findings. First, high-quality content and extensive communication influence are identified as dual necessary conditions, resolving the theoretical tension between value-centric and reach-centric logics. Second, we uncover two equifinal sufficient configurations built upon this shared foundation: a "Community Builder" strategy that combines the core conditions with high affinity, and a "Flagship" strategy that combines them with demonstrative leadership. This study develops and validates a configurational model of digital success, offering a nuanced theoretical alternative to correlational approaches and providing a strategic map for practitioners navigating the complexities of digital communication.

Keywords: Fuzzy-Set Qualitative Comparative Analysis (Fsqca); Configuration Theory; Higher Education; Digital Strategy; Ideological and Political Education (Ipe); Wechat Accounts; Chinese University

1. Introduction

In the contemporary digital landscape, social media is no longer a peripheral marketing tool for universities but has become a central arena for institutional branding, stakeholder engagement, and the fulfillment of core educational missions (Capriotti et al., 2024; Rutter et al., 2016). Within China, this digital transformation is especially critical for the practice of Ideological and Political

Education (IPE), a foundational component of the national higher education system (Yu & Yao, 2024) University IPE WeChat accounts have thus emerged at a critical intersection, tasked with the complex challenge of translating institutional values into engaging content for a digitally native generation that exhibits distinct patterns of media consumption and identity expression, particularly on mobile and video-centric platforms (Lian, 2023; Zhou et al., 2024). However, many of these accounts struggle to translate effort into effect. Existing studies and practical observations point to widespread challenges such as content homogenization, superficial engagement, and a polarization of communication effects, where official announcements fail to resonate with the target audience (Arora et al., 2022; Ngai et al., 2020). The success in this digital realm directly impacts an institution's legitimacy and its ability to fulfill its fundamental purpose.

Faced with this challenge, university administrators and scholars alike have sought to identify the key drivers of success (Shu, 2023; Yang et al., 2023). Existing research has made valuable contributions by consistently pointing to a series of critical factors, such as high-quality content, extensive communication reach, and high levels of user affinity (Lian, 2023; Yang et al., 2023; Zhang, 2023). However, the vast majority of this research relies on variable-centered, linear-additive models such as regression analysis (Chen, 2024; Shen et al., 2019; Sun et al., 2020; Yang et al., 2023), which assess the isolated "net effect" of each factor. This dominant approach is often criticized for being fundamentally ill-equipped to capture the complex realities of strategy, as it assumes factors work in isolation and are simply additive (Fiss, 2011). By searching for a "one-size-fits-all" solution, it generates a fragmented and often confusing "checklist" of best practices, leaving a critical research gap that how can we understand these crucial factors combine synergistically to produce high impact.

Without a coherent map of successful strategic combinations, university managers are left to navigate the complex digital environment with inadequate tools, often leading to wasted resources on fragmented initiatives that fail to deliver a synergistic impact (Luan et al., 2020). To address this critical gap, this study investigates the causal complexity behind high-performing IPE WeChat accounts on consideration of Chinese background. We move beyond simplistic linear assumptions to ask a more nuanced, holistic question: What configurations of platform characteristics lead to high impact for university IPE WeChat accounts? To answer this central question, we employ fuzzy-set Qualitative Comparative Analysis (fsQCA), a set-theoretic method specifically designed to identify complex causal recipes and uncover conditions of equifinality, the existence of multiple pathways to the same outcome(Fiss, 2011; Ragin, 2008; Schneider & Wagemann, 2012). This configurational approach is ideal for understanding how conditions combine synergistically, and it allows us to investigate three specific sub-questions: (1) Are there multiple, equally effective pathways to achieving high impact? (2) Are any individual characteristics necessary for success? and (3) Is the explanation for success the symmetrical inverse of the explanation for failure?

This study offers significant value to both theory and practice. Theoretically, it challenges the prevailing linear models in digital communication research by introducing and empirically testing a configurational approach. Practically, the findings will provide university administrators and communication managers not with another simplistic checklist, but with a "strategic map" of

distinct, actionable, and context-dependent pathways for enhancing the impact of their digital IPE efforts.

2. Literature Review

This chapter provides the theoretical foundation for the study. It first reviews existing research on IPE WeChat accounts to identify a critical gap, then synthesizes broader theories of communication and institutional legitimacy to develop the configurational model.

2.1. Studies on IPE WeChat Accounts and Their Limitations

The strategic importance of IPE WeChat accounts has spurred a growing body of research (Lian, 2023; Shu, 2023). Scholars have made valuable contributions by identifying a series of factors critical to their success. These include, for example, the necessity of high-quality, mission-aligned content (Zhang, 2023) , the importance of extensive communication reach (Yang et al., 2023) , and the role of high-level user affinity and engagement (Lian, 2023). However, the vast majority of this research relies on variable-centered, linear-additive models, such as regression analysis, to assess the isolated "net effect" of each factor (Chen, 2024; Sun et al., 2020; Yang et al., 2023). This dominant approach, while useful, often generates a fragmented and sometimes confusing "checklist" of best practices. It is fundamentally ill-equipped to capture the complex realities of strategy, where factors work in synergistic combination rather than in isolation (Fiss, 2011). A critical research gap thus remains: we know what factors matter, but we do not know how these crucial factors combine synergistically to produce high impact. Existing research, with its linear assumptions, cannot adequately explain why some accounts with seemingly "good" content fail, while others succeed through different strategic profiles. This study addresses this gap by moving beyond a "net effects" logic to ask a holistic, configurational question: What combinations of platform characteristics lead to high impact for university IPE WeChat accounts?

2.2. The Digital Transformation of University Communication and IPE

The ubiquity of social media has reshaped higher education by enhancing communication, collaboration, and access to resources (Aleksandrova & Parusheva, 2019). Platforms like WeChat in China have evolved from peripheral channels into central arenas for institutional communication, profoundly altering how universities manage their reputation, engage with stakeholders, and fulfill their educational missions (Pang, 2020; Zhang, 2023; Zhang et al., 2024). This digital transformation is especially critical for IPE, a core function of Chinese universities. The traditional top-down, classroom-based model of Ideological and Political Education faces challenges in engaging a digitally native student body accustomed to interactive and user-centric media (Chu, 2023). Consequently, university IPE WeChat accounts have become strategic assets, tasked not only with disseminating information but with actively shaping values and fostering community in a complex and noisy digital ecosystem (Lian, 2023; Zhou et al., 2024). Therefore, understanding what constitutes "effectiveness" in this new context is a pressing theoretical and practical concern.

2.3. The Duality of Digital Influence: Mass Communication vs. Community Engagement

To understand effectiveness on social media platforms, we must first consider the competing logics of digital influence rooted in communication theory. The first logic, stemming from classic mass communication theories, equates influence with reach and exposure (Katz & Lazarsfeld, 1955). From this perspective, an effective platform is one that maximizes its audience size, viewership, and information dissemination capacity (Kandhway & Kuri, 2016; Purtle et al., 2020). This "broadcast" model prioritizes quantitative metrics and is built on the premise that a larger audience translates to greater potential impact (Zhang et al., 2020). The second, contrasting logic emphasizes community engagement and resonance (Brunton et al., 2017; Weitz, 2023). Drawing from Uses and Gratifications Theory (Rubin, 2009), this view posits that users are not passive recipients but active agents who use media to fulfill social and emotional needs. Therefore, influence is not just about reaching people but about resonating with them (Liu et al., 2023). It is achieved by building relationships, fostering dialogue, and creating a sense of belonging within an online community (Overstreet, 2020). This "community" model prioritizes qualitative engagement, loyalty, and the strength of the user-platform bond (Kumar et al., 2022; Zheng et al., 2015). A critical tension exists that should a platform prioritize broadcasting to the many or building deep relationships with a dedicated community? Successfully navigating this duality is a central challenge for digital strategists.

2.4. The Institutional Imperative: Legitimacy as the Ultimate Goal

While communication theories explain the mechanisms of influence, Institutional Theory explains the motivation behind a university's digital strategy (Hinings et al., 2018; Valdés-León et al., 2021). As organizations, universities operate within a complex institutional field and are driven by a powerful need for legitimacy, the perception that their actions are appropriate and desirable within their social and political context (DiMaggio & Powell, 1983; Scott, 2013). An IPE WeChat account is a primary tool for signaling and securing this legitimacy. This institutional imperative manifests in several ways. Firstly, it demands a high degree of responsiveness to key directives, a form of normative and coercive isomorphism where the university aligns its messaging with the expectations of key stakeholders like the government and the Ministry of Education to maintain its standing as a legitimate public institution (Dias et al., 2021; Salmi, 2007; Varughese, 2017). Secondly, universities engage in competitive strategies to enhance their status, such as striving for demonstrative leadership in digital innovation to be seen as a sector leader (Isaeva et al., 2024; Kiriliuk & Zakharova, 2024). Crucially, the ultimate foundation of an educational institution's legitimacy is the substance of its mission. Therefore, the ideological and educational content it produces is not merely informational but is the very embodiment of its claim to authority and purpose (Verhoeven & Verbruggen, 2024). High-quality, mission-aligned content is the core resource that underpins all other legitimacy seeking activities on the platform. Thus, the institutional context compels universities to manage a portfolio of actions, which includes responsiveness, leadership, and substantive content, to secure and enhance their legitimacy.

2.5. A Configurational Synthesis: Multiple Pathways to Legitimacy and Impact

From the above, the core logic of institutional social accounts includes the communication duality of reach versus resonance, and the institutional imperative for legitimacy through content, leadership, and responsiveness. A simple linear model (Chen, 2024; Shen et al., 2019; Sun et al., 2020; Yang, 2023) is inadequate to capture how these forces combine. Therefore, this study adopts Configuration Theory as its synthesizing framework (Fiss, 2011; Meyer et al., 1993). Configuration theory posits that effectiveness arises not from individual factors but from the holistic and synergistic interplay among them, and that multiple, different combinations of factors or configurations, can lead to the same successful outcome (Andrews et al., 2016; Puppatz et al., 2023). This theoretical lens allows us to integrate the previously discussed dimensions into a cohesive model. So we propose that high-impact IPE WeChat accounts achieve success through different, and equifinal configurations of communication logics and institutional strategies.

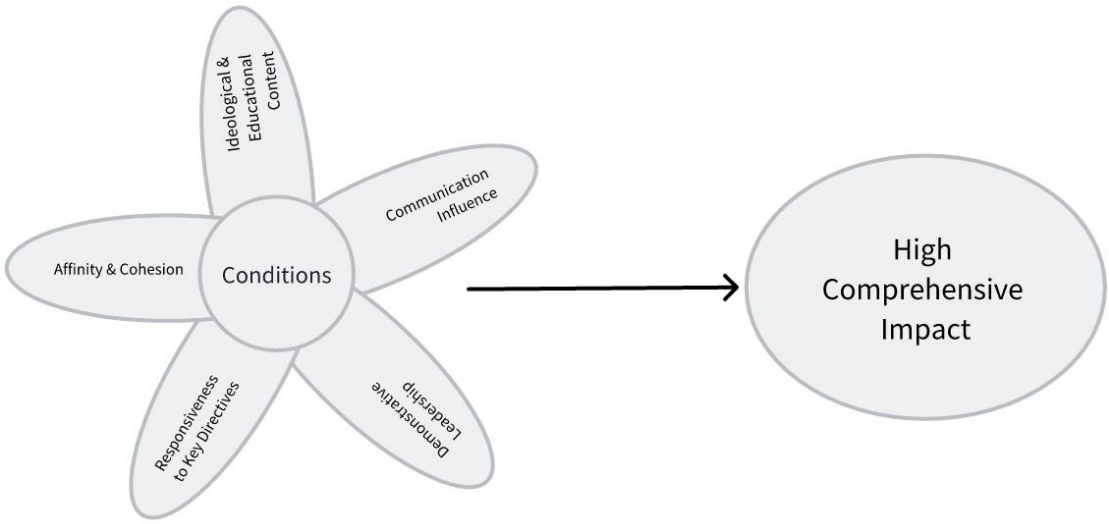

Figure 1. Conceptual Model

Figure 1 visually presents our conceptual model, which integrates the theoretical streams discussed above. The model identifies five key conditions posited to combine to produce the outcome. Based on the communication duality, we derive two conditions: Communication Influence (CI) , representing the “mass communication” logic of reach, and Affinity & Cohesion (AC), representing the “community engagement” logic of resonance. From the institutional imperatives (Section 2.4), we derive three conditions: the foundational Ideological & Educational Content (IEC), which embodies the IPE mission; Demonstrative Leadership (DL) , reflecting institutional status-seeking; and Responsiveness to Key Directives (RKD), reflecting organizational alignment. The configurational argument is that High Comprehensive Impact is not achieved by maximizing any single condition, but rather through distinct, synergistic combinations of these five conditions.

In conclusion, based on this configurational synthesis, this study operationalizes the key elements of effectiveness by drawing from these theoretical streams. From the communication logics, we derive the conditions of Communication Influence and Affinity and Cohesion. Concurrently, from the institutional imperatives of legitimacy, we derive the conditions of Responsiveness to Key Directives, Demonstrative Leadership, and the foundational element of

Ideological and Educational Content. These five conditions constitute the building blocks of the configurations we seek to uncover through fsQCA (as shown in the Conceptual Model of Figure 1), allowing us to explore the multiple, complex pathways to achieving high impact in the digital landscape of Ideological and Political Education.

3. Methodology

This section outlines the fsQCA methodology. It covers the case selection and data source, the operationalization of the study's conditions, and the multi-step procedures used for data analysis.

3.1. Research Design

This study adopts the method of fsQCA to answer the research questions that how could these crucial factors combined synergistically to produce high impact. Unlike conventional quantitative methods such as regression analysis, which focus on the net effect of independent variables, fsQCA is a set-theoretic approach rooted in configurational theory (Fiss, 2011; Ragin, 2008). It is exceptionally well-suited for this study for three reasons. First, fsQCA is designed to unravel causal complexity, identifying conjunctural causation, where conditions are only effective in combination (Ding, 2022; Liu et al., 2023); equifinality, where multiple distinct causal pathways can lead to the same outcome (Cicchetti & Rogosch, 1996; Wolfschwenger & Young, 2021); and causal asymmetry, where the explanations for the presence of an outcome can be different from the explanations for its absence (Schneider & Wagemann, 2012). This aligns with our theoretical assumption that there is no single best way to achieve high impact. Second, fsQCA is ideal for small- to medium-N research designs (N=30 in our case), where traditional statistical methods lack sufficient power and case-oriented qualitative methods lack generalizability (Trueb, 2013). It bridges the gap between these two approaches by enabling a systematic, cross-case analysis while preserving the holistic integrity of each case (Douglas et al., 2020; Kumar et al., 2022). Third, by using fuzzy sets, fsQCA moves beyond simple binary (present/absent) classifications and allows for the calibration of partial membership of cases in sets (Ragin, 2006). This accommodates the nuanced reality that universities possess characteristics (e.g., “communication influence” or “affinity”) to varying degrees, providing a more fine-grained analysis.

3.2. Case Selection and Data Source

The cases for this study are the 30 top-performing Chinese universities as listed in the April 2025 “Ranking of Human Power in Ideological and Political Education”. This ranking constitutes an information-rich sample of elite institutions that have demonstrated success in digital IPE, making them ideal for an inquiry into the configurations driving high performance. The data was sourced from the China National Ideological and Political Work Net (www.sizhengwang.cn), an official and authoritative platform supervised by the Ministry of Education of the People’s Republic of China. This monthly report assesses IPE accounts using a comprehensive index that includes multiple indicators tracking, for example, content quality, communication reach, engagement, and innovation. These indicators align closely with our chosen conditions (e.g., IEC, CI, AC, DL), thus ensuring the data's high relevance and construct validity for this study. The specific dataset is from the monthly report covering the period from April 1st, 2025, to April 30th,

2025. This official source ensures the credibility and relevance of the data for assessing IPE effectiveness in China's higher education context.

Table 1. The Original Data Collected

Case Number	IEC	CI	DL	RKD	AC	HCI
1	2.58	2.88	0.98	0.89	0.9	823.46
2	2.56	2.88	0.98	0.9	0.92	822.79
3	2.49	2.88	0.9	0.89	0.93	809.38
4	2.54	2.84	0.77	0.9	0.93	798.96
5	2.43	2.82	0.95	0.87	0.9	796.76
6	2.44	2.85	0.9	0.87	0.88	793.71
7	2.51	2.8	0.82	0.89	0.91	793.22
8	2.43	2.81	0.89	0.87	0.91	790.3
9	2.44	2.78	0.92	0.87	0.88	789.59
10	2.42	2.81	0.88	0.88	0.91	789.51
11	2.4	2.8	0.93	0.85	0.9	788.77
12	2.51	2.8	0.76	0.88	0.92	786.26
13	2.51	2.78	0.74	0.89	0.92	784.11
14	2.41	2.78	0.88	0.87	0.9	783.96
15	2.43	2.75	0.87	0.87	0.91	783.52
16	2.39	2.74	0.96	0.86	0.88	783.07
17	2.4	2.77	0.86	0.85	0.87	775.8
18	2.36	2.75	0.87	0.85	0.89	773.51
19	2.37	2.74	0.88	0.86	0.86	770.82
20	2.34	2.73	0.89	0.85	0.88	769.01
21	2.34	2.69	0.85	0.86	0.93	767.03
22	2.31	2.73	0.89	0.85	0.89	765.9
23	2.32	2.72	0.89	0.84	0.85	762.01
24	2.33	2.69	0.87	0.85	0.89	761.65
25	2.29	2.75	0.86	0.84	0.87	760.96
26	2.3	2.7	0.89	0.84	0.87	760.26
27	2.26	2.7	0.91	0.84	0.89	759.93
28	2.34	2.72	0.76	0.86	0.89	757.09
29	2.24	2.72	0.85	0.84	0.9	755.17
30	2.26	2.68	0.87	0.84	0.89	754.88

3.3. Description of Conditions and Outcome

Our configurational model incorporates one outcome, High Comprehensive Impact (HCI), and five causal conditions. These conditions are derived from our conceptual model (Fugure 1) and are chosen to holistically capture the key strategic dimensions of content quality, communication strategy, and institutional positioning. The detailed definition for each variable is provided in Table 2.

Table 2. Definition of Outcome and Causal Conditions

Abbreviation	Variable Name	Description
HCI	High Comprehensive Impact	(Outcome) The overall performance and effectiveness of the university's IPE WeChat account, measured by the comprehensive score in the ranking.
IEC	Ideological & Educational Content	(Condition) The quality and depth of content in promoting core values and mainstream ideology.
CI	Communication Influence	(Condition) The account’s reach and engagement, reflected in metrics like readership and user base size.
DL	Demonstrative Leadership	(Condition) The account's role as an industry benchmark in innovation and agenda-setting.
RKD	Responsiveness to Key Directives	(Condition) The effectiveness in communicating and interpreting major policies and university-level initiatives.
AC	Affinity & Cohesion	(Condition) The ability to build an emotional connection and a sense of community with the audience.

3.4. Analysis Approach

The data analysis followed a structured, multi-step procedure using the fsQCA 4.0 software package.

Calibration. Calibration is the process of transforming the raw scores for each variable into fuzzy-set membership scores, which range from 0 (full non-membership) to 1 (full membership). We employed the direct calibration method (Ragin, 2008), a standard procedure that uses theoretical knowledge and empirical data to define three qualitative anchors: the threshold for full membership (fuzzy score of 0.95), the threshold for full non-membership (fuzzy score of 0.05), and the crossover point (fuzzy score of 0.5), indicating maximum ambiguity. Given the absence of established external standards, we determined these anchor points based on the percentile distribution of our sample (N=30). Specifically, we set the thresholds for full membership (fuzzy score of 0.95), the crossover point (fuzzy score of 0.5), and full non-membership (fuzzy score of 0.05) to correspond with the 75th percentile, 50th percentile (median), and 25th percentile of the

raw scores, respectively. The specific values for these anchors for all conditions are detailed in Table 3.

Table 3. Descriptive Statistics for Dependent and Independent Variables

		IEC	CI	DL	RKD	AC	HCI
Mean		2.398	2.770	0.876	.864	0.896	780.380
Variance		0.009	0.003	0.004	0.000	0.000	353.804
Minimum		2.240	2.680	0.740	0.840	0.850	754.880
Maximum		2.580	2.880	0.980	0.900	0.930	823.460
Percentile	25	2.328	2.720	0.858	0.850	0.880	761.920
	50	2.400	2.760	0.880	0.860	0.895	783.295
	75	2.453	2.810	0.903	0.880	0.910	791.030

Analysis of necessary conditions. The study conducted an analysis of necessary conditions to test whether any single condition (or its negation) was a prerequisite for the outcome by examining if its consistency score exceeded the 0.90 threshold (Schneider & Wagemann, 2012).

Construction of truth table. Following calibration, we constructed a truth table listing all 2^5=32 logically possible combinations of the five conditions, and cases were assigned to analyse all possible combinations of causal conditions. The truth table was structured in binary form to represent the presence or absence of each condition (Ragin, 2008). To ensure analytical rigour and empirical relevance, we refined the truth table using established criteria: a minimum frequency threshold of one case and a consistency threshold of 0.80 (Li, 2018). This refinement step ensured that only empirically meaningful configurations were retained for subsequent analysis, while maintaining the methodological standards of fsQCA.

Configuration analysis. With the refined truth table, we conducted a configuration analysis using the Quine-McCluskey algorithm and counterfactual analysis (Fiss, 2011). This analytical process systematically identified consistent causal combinations while distinguishing between core and peripheral conditions (S. Chen et al., 2021). Through Boolean minimization, the analysis revealed how different conditions interrelate within successful configurations to produce high dissemination effectiveness. FsQCA software was used to perform the analysis while ensuring both analytical rigor and theoretical relevance, thus allowing for a thorough exploration of the necessary and sufficient conditions for achieving high comprehensive impact.

Predictive validity analysis. To validate the robustness of the configurational solutions, we conducted a predictive validity analysis (Woodside, 2014). The dataset was randomly split into two equal subsamples using SPSS. FsQCA was applied to both subsamples using the same analytical parameters as in the main study (Pappas et al., 2016). The resulting configurational

model was then tested against the holdout subsample to assess its predictive capability. This cross-validation approach enabled us to evaluate the consistency and coverage levels across the different datasets, thereby confirming the generalisability of our findings (Olya & Altinay, 2016).

Sensitivity analysis. To ensure the robustness of our findings, we performed a sensitivity analysis by examining alternative calibration specifications (Fiss, 2011). Specifically, we systematically adjusted the initial calibration anchors (upper quartile, median, and lower quartile values) by ±10%. This analysis assessed the stability of our configurational solutions and verified that the results remained consistent across different calibration specifications, ensuring that the findings were not dependent on specific calibration thresholds.

Post hoc analysis. As a supplementary post hoc analysis, we integrated the fsQCA solutions into a Tobit regression framework (Fiss, 2011) to explore the marginal effects of the identified configurations on comprehensive impact. In this analysis, comprehensive impact was the dependent variable, and the configurational solutions from fsQCA were treated as independent variables. While the regression analysis was not intended to replace or confirm the fsQCA solutions, it provided additional insights into the relative importance of each path. This complementary analysis enhanced our understanding of the causal relationships identified through fsQCA, offering supplementary perspectives on the marginal effects of the configurations.

4. Results

This section presents the findings from the fsQCA analysis. It reports on the necessary conditions, details the two sufficient configurations for high impact, and summarizes the results of the robustness checks.

4.1. Results of Necessary Conditions

A necessary condition is one that must be present for an outcome to occur. In set-theoretic terms, the outcome set is a subset of the condition set. We followed the standard convention of setting a consistency score threshold of 0.90 to identify a necessary condition (Schneider & Wagemann, 2012).

As shown in Table 4, two conditions surpass the 0.90 consistency threshold for the presence of high comprehensive impact. High-quality Ideological & Educational Content (IEC) has a consistency score of 0.923, and high Communication Influence (CI) has a consistency score of 0.934. This indicates that both strong content and wide reach function as foundational, necessary conditions for achieving high comprehensive impact among China's top universities. No other single condition qualifies as necessary for the outcome's presence.

Symmetrically, the analysis for the absence of high impact (~HCI) reveals that the absence of content (~IEC, consistency = 0.902) and the absence of influence (~CI, consistency = 0.905) are necessary conditions for non-high impact. This further strengthens the finding that IEC and CI are the fundamental prerequisites for entering the ranks of high-performing accounts.

Table 4. Analysis of Necessary Conditions

Condition	HCI		~HCI	
	Consistency	**Coverage**	**Consistency**	**Coverage**
IEC	0.923	0.902	0.218	0.219
~IEC	0.2	0.199	0.902	0.924
CI	0.934	0.905	0.232	0.232
~CI	0.207	0.207	0.905	0.934
DL	0.678	0.653	0.445	0.442
~DL	0.421	0.424	0.651	0.676
RKD	0.867	0.837	0.263	0.262
~RKD	0.236	0.237	0.836	0.866
AC	0.757	0.732	0.34	0.339
~AC	0.317	0.318	0.731	0.756
Note: A consistency score > 0.90 is required to classify a condition as necessary.				

4.2. Results of Configurations Analysis

Following the necessity analysis, we performed a sufficiency analysis to identify the configurations leading to high comprehensive impact. Using a frequency threshold of 1 and a consistency cutoff of 0.873, the analysis yielded a final solution with a high overall consistency of 0.928 and a strong overall coverage of 0.846 (as shown in Table 5). This indicates that the model reliably explains the vast majority of high-impact cases. The analysis revealed two distinct, equifinal pathways to success.

The initial software-generated intermediate solution yielded two pathways. However, a core principle of set-theoretic research is that identified necessary conditions should be present in all sufficient configurations. Our necessity analysis established that both Ideological & Educational Content (IEC) and Communication Influence (CI) are necessary conditions for high impact. While the initial solution's paths both contained IEC, the first path did not include CI.

Therefore, to enhance the theoretical coherence and rigor of our findings, we constrained the final solution based on these results. We manually incorporated Communication Influence (CI) as a required condition into the first pathway, ensuring that both final configurations presented are

built upon the foundation of all identified necessary conditions. This theoretically informed adjustment yielded a final solution with an overall consistency of 0.921 and coverage of 0.835.

Table 5. Configurations for Achieving High Comprehensive Impact (HCI)

Conditions	**Content-Affinity Driven Pathway**	**Leadership-Integration Driven Pathway**
Ideological & Educational Content	●	●
Communication Influence	●	●
Demonstrative Leadership	⊗	⊗
Responsiveness to Key Directives	⊗	⊗
Affinity & Cohesion	⊗	
Consistency	0.854	0.997
Raw Coverage	0.372	0.590
Unique Coverage	0.257	0.475
Solution Coverage	0.846	
Solution Consistency	0.928	
Note: Note: ● (large solid circle): Core condition present. ⊗ (large crossed-out circle): Peripheral condition present. ⊗ (small crossed-out circle): Peripheral condition absent. Blank spaces indicate a "don't care" condition.		

The first pathway is the "Content-Affinity Driven" configuration. In its refined form, this path is defined by the core presence of both necessary conditions—Ideological & Educational Content and Communication Influence. This dual core is then combined with two peripheral conditions: the presence of high Affinity & Cohesion and the absence of Demonstrative Leadership. This pathway, representing a strategy of deep engagement built upon a strong content and communication foundation, shows a consistency of 0.841 and uniquely explains 0.243 of the high-impact cases.

The second pathway is the "Leadership-Integration Driven" configuration. This path also features the core presence of Ideological & Educational Content and Communication Influence. It combines this core foundation with the peripheral presence of Demonstrative Leadership. This configuration, which represents a strategy of leveraging broad influence and industry status, is highly consistent at 0.997 and is the empirically more dominant path, uniquely explaining 0.475 of successful outcomes.

This more rigorous analysis presents a clear and powerful finding. It confirms the existence of two distinct pathways to success, but more importantly, it demonstrates that both pathways are built upon the same dual-core foundation: high-quality content combined with strong communication reach. The strategic differentiation between high-performing institutions,

therefore, occurs in the choice of peripheral elements, either a focus on community-building or on institutional leadership, that are arranged around this shared, non-negotiable core.

4.3. Validation and Post-hoc Analysis Results

The sensitivity analysis, which utilized a more stringent 95/50/5 percentile calibration, yielded results substantially consistent with the main analysis. While minor variations in consistency and coverage scores were observed, the core solution identifying the "Content-Affinity Driven" path and the "Leadership-Integration Driven" path remained stable. This confirms that our findings are robust and not dependent on a specific calibration scheme.

The configurations derived from Sample 1 were tested using the data from Sample 2, yielding comparable levels of consistency and coverage (see Figures 2 and 3). These results support the assumption that the proposed configuration model demonstrates high predictive validity.

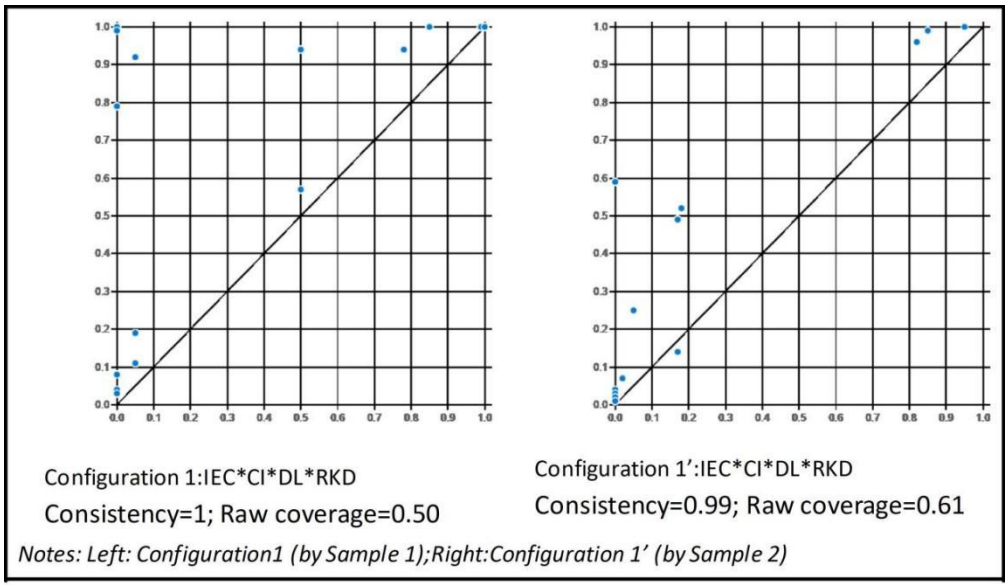

Figure 2. XY scatterplots of High Comprehensive Impact in Configurations 1 and 1'

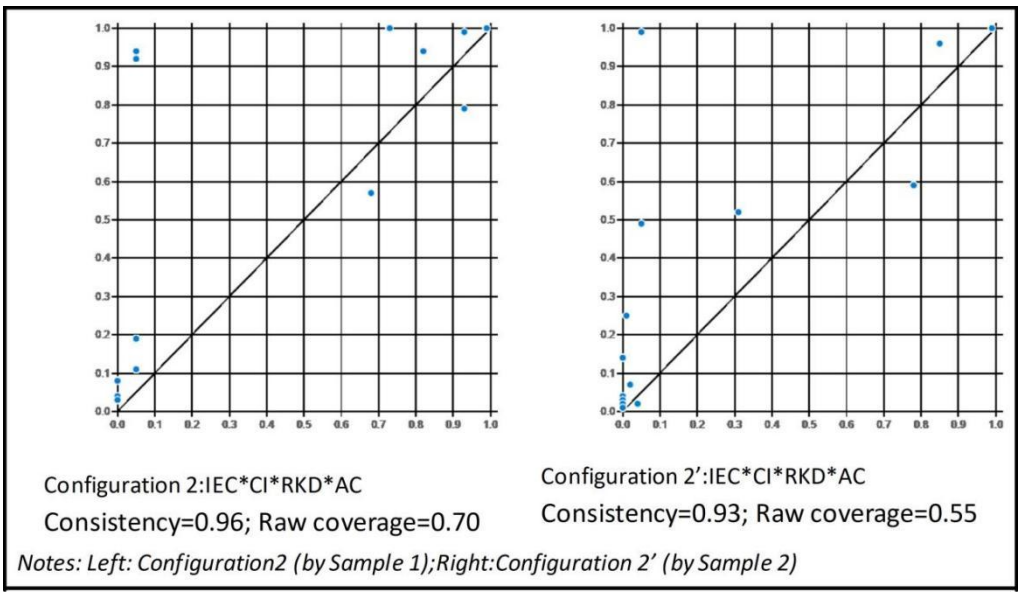

Figure 3. XY scatterplots of High Comprehensive Impact in Configurations 2 and 2'

The results of the Tobit regression provide strong, unambiguous support for our configurational findings (as shown in Table 6). As the analysis shows, the membership score in both the Content-Affinity Path (Coefficient = 0.758, $p < 0.001$) and the Leadership-Integration Path (Coefficient = 1.166, $p < 0.001$) has a positive and highly statistically significant effect on the comprehensive impact score.

This quantitative analysis confirms that both of the configurations identified through fsQCA are substantively and significantly linked to achieving higher performance. The positive coefficients indicate that a university's closer conformity to either of the two successful strategic recipes is associated with a greater marginal increase in its final impact score. Furthermore, the analysis indicates that the coefficient for second path is larger than that for first one, suggesting that membership in the "Leadership-Integration Driven" path is associated with a greater marginal effect, a finding that is consistent with its higher raw coverage in the primary fsQCA analysis. This post-hoc test therefore provides robust correlational evidence that validates and complements our main set-theoretic conclusions.

Table 6. Post-hoc Tobit Regression on Comprehensive Impact Score

Variable	Coefficient	Std. Error	z-Statistic	Prob.
Content-Affinity Path	0.758361	0.098635	7.688587	0.0000
Leadership-Integration Path	1.165619	0.082905	14.05974	0.0000

5. Discussion

This section interprets the study's findings. It discusses their theoretical implications in relation to existing literature and outlines the practical applications for university managers.

5.1. Key Findings

Our analysis yields three principal findings that together form a configurational map of digital IPE success. First, the necessity analysis established that both high-quality Ideological & Educational Content and high Communication Influence are foundational prerequisites for achieving a high level of impact. This indicates that to even be in the running for excellence, an account must both produce valuable content and ensure it reaches a substantial audience. Second, and central to this study, is the discovery of the two successful strategic recipes. The "Content-Affinity Driven" path demonstrates a strategy of deep community engagement, while the "Leadership-Integration Driven" path showcases a strategy of leveraging institutional prestige and mass communication. The existence of these two viable pathways provides clear empirical evidence for equifinality. Third, our analysis revealed a clear dual-core structure within both configurations, with Ideological & Educational Content and Communication Influence serving as the non-negotiable foundation, while other elements such as Affinity & Cohesion and Demonstrative Leadership function as peripheral, yet defining, components of each distinct strategy.

5.2. Theoretical Implications

These empirical findings have significant theoretical implications. First and foremost, this study provides a direct empirical answer to the research gap identified in Section 2.1. Instead of a fragmented “checklist” of success factors, our results support the core tenets of Configuration Theory (Meyer et al., 1993), demonstrating that impact arises from the holistic interplay of strategic elements (Fiss, 2011; Puppatz et al., 2023). By identifying two equifinal pathways, our findings empirically challenge the “one-size-fits-al” logic inherent in the variable-centered, linear-additive models that are prevalent in IPE communication research (Chen, 2024; Sun et al., 2020). This marks a theoretical shift from asking “what factors matter” to “how factors combine to matter” in achieving digital success.

The two pathways themselves can be interpreted as distinct resolutions to the theoretical "duality of digital influence" discussed earlier. The "Content-Affinity Driven" path empirically validates the community engagement logic (Brunton et al., 2017; Weitz, 2023). It shows how platforms can succeed not simply by maximizing reach, but by resonating deeply with users, satisfying their social and emotional needs as described by Uses and Gratifications Theory (Rubin, 2009), and thereby fostering a strong user-platform bond. In contrast, the “Leadership-Integration Driven” path exemplifies the powerful synergy between the classic mass communication logic of maximizing audience size and exposure (Kandhway & Kuri, 2016; Katz & Lazarsfeld, 1955) and the institutional imperative for status-seeking. This aligns perfectly with Institutional Theory's view of organizations proactively shaping their environment to enhance their prestige and legitimacy (DiMaggio & Powell, 1983; Isaeva et al., 2024).

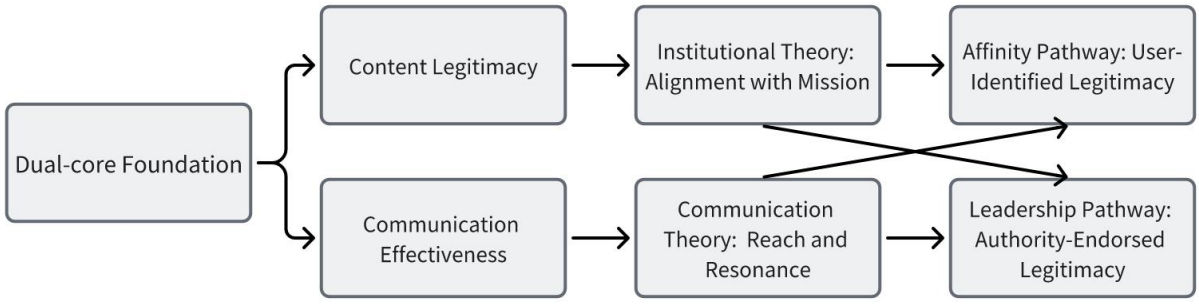

Figure 4. Dual-Core Theoretical Integration Framework for Digital IPE Success

Furthermore, our findings connect to and extend other key communication theories. The "Content-Affinity Driven" path provides strong empirical support for User-Generated Content (UGC) Theory. This configuration's reliance on high “Affinity & Cohesion” demonstrates that success is not merely broadcast-driven, but is co-created by fostering a community where users feel a sense of belonging and are motivated to engage, comment, and share, aligning with the "Uses and Gratifications" perspective (Rubin, 2009). In parallel, the model aligns with Media Convergence Theory. WeChat itself is a converged platform, blending mass communication (CI), social networking (AC), and institutional branding (DL) .The findings empirically show that high impact does not arise from mastering one of these functions in isolation, but from a strategic

configuration of these converged elements, offering two distinct, successful models for navigating this complexity.

Perhaps the most crucial theoretical contribution is the synthesis offered by our finding of a dual-core structure (as shown in Figure 4). It suggests that while universities can choose between different peripheral strategies (community vs. leadership), they cannot escape the foundational requirements of both producing substantive content and achieving significant communication influence. The centrality of Ideological & Educational Content confirms that the institutional need to provide mission-aligned substance as the basis of legitimacy is paramount (Scott, 2013; Verhoeven & Verbruggen, 2024). This grounds the entire digital strategy in the university's fundamental purpose, a point underscored by recent work on the importance of digital IPE in engaging contemporary students (Chu, 2023; Lian, 2023).

5.3. Practical and Managerial Implications

Our findings offer a clear and actionable "strategic map" for university administrators, moving beyond the often confusing and fragmented advice that can lead to wasted resources (Luan et al., 2020). The most fundamental implication is the need for a managerial mindset shift: from searching for isolated "best practices" to adopting a holistic, configurational perspective. Rather than asking "Should we focus on content or on promotion?", leaders should ask "Which complete recipe of synergistic actions is most coherent with our institution's unique identity and resources?". This approach replaces a simple checklist with a more sophisticated model of strategic alignment.

The analysis then provides a clear, two-step sequence for strategic planning. The first step is to secure the non-negotiable foundation. Our results are unequivocal that all successful strategies begin with the same dual-core foundation: high-quality Ideological and Educational Content combined with significant Communication Influence. This provides a clear directive for resource allocation. It means universities must simultaneously invest in both the "substance" of their message (content research, creation, and quality control) and the "delivery" of that message (platform promotion, audience growth, technical optimization). Excelling in only one of these two core areas is an insufficient condition for achieving high impact.

Once this dual-core foundation is in place, managers face a clear and evidence-based strategic choice between two proven pathways. The "Community Builder" path is a strategy of depth over breadth, focusing on authenticity and fostering deep, affective engagement to build a loyal and interactive community (Zhang, 2023). This approach is potent for any institution aiming to strengthen its unique cultural identity and build strong stakeholder relationships. Alternatively, the "Flagship" path is a strategy of leadership and scale, which involves leveraging institutional status to lead national discourse and set industry benchmarks (Kiriliuk & Zakharova, 2024). This path is best suited for well-resourced, top-tier universities aiming to project their influence and solidify their elite position. The power of these findings lies in validating both approaches, allowing for more deliberate and synergistic strategic planning that aligns with an institution's specific capabilities and goals.

6. Conclusion, Limitation and Future Research

This final section summarizes the study's main contributions, acknowledges its limitations, and proposes directions for future research.

6.1. Conclusion

This study sought to move beyond simplistic, single-factor explanations of success in the digital realm of Ideological and Political Education. By employing a configurational approach, we have provided a more holistic and causally complex understanding of how China's universities achieve high impact with their IPE WeChat accounts. Our central contribution is the empirical identification of two distinct and equifinal pathways to success. The "Content-Affinity Driven" path leverages deep community engagement, and the "Leadership-Integration Driven" path relies on the synergy of massive reach and institutional prestige. Crucially, we found that both of these successful strategies are built upon the same non-negotiable dual-core foundation of high-quality Ideological & Educational Content and strong Communication Influence. In doing so, this research offers a nuanced configurational theory of digital success that synthesizes communication and institutional logics, and provides a context-dependent strategic map for university administrators navigating the complexities of digital communication.

6.2. Limitation and Future Research

Like all research, this study defines clear boundaries that, in turn, illuminate a clear and productive agenda for future inquiry. Our deliberate focus on elite, top-performing universities was crucial for understanding the configurations that produce excellence, yet it naturally raises questions about generalizability. Future research should therefore apply this configurational model to a more diverse sample of mid-tier, regional, and specialized institutions to discover whether the pathways identified here are universal, or if distinct configurations for success emerge under different resource constraints.Furthermore, the study's cross-sectional design provides a powerful snapshot of successful strategies but does not capture their evolution. A compelling avenue for future work lies in longitudinal research, employing time-series Qualitative Comparative Analysis (QCA) to map how universities develop, maintain, and perhaps transition between these strategic configurations over time. Third, our study carries the inherent methodological limitations of fsQCA. The operationalization of conditions and, most notably, the calibration of anchor points, involves a degree of researcher judgment. While we defended our choices based on the data's percentile distribution (Table 3) and confirmed their robustness through sensitivity analysis , this potential for subjective bias remains. Future research could explore alternative calibration strategies or different operationalizations of the conditions to further test the stability of these configurations. Finally, while our analysis identifies what successful recipes contain, it invites deeper inquiry into how they are executed. We strongly advocate for future research using in-depth qualitative case studies of universities that are archetypal examples of each path. Such work could illuminate the micro level managerial decision-making, team dynamics, and specific content practices that bring these configurations to life, and then building a richer and more dynamic theory of digital IPE success.

Author Contributions:

Conceptualization, S.W.; methodology, S.W.; software, S.W.; validation, S.W.; formal analysis, S.W.; investigation, S.W.; resources, S.W.; data curation, S.W.; writing—original draft preparation, S.W.; writing—review and editing, S.W.; visualization, S.W.; supervision, S.W.; project administration, S.W.; funding acquisition, S.W. All authors have read and agreed to the published version of the manuscript.

Funding Statement:

This work was jointly supported by the Special Project of Shanghai Municipal Philosophy and Social Sciences Planning (2025VSZ023), the Shanghai Education Science Research Project Grant (2024ZSD032), the Shanghai Association of Higher Education Grant (2QYB24027), and the Shanghai Arts Science Planning Project Grant (YB2024-H-035).

Institutional Review Board Statement:

Not applicable.

Informed Consent Statement:

Not applicable.

Data Availability Statement:

Not applicable.

Acknowledgments:

Not applicable.

Conflict of Interest:

The author declared no potential conflicts of interest with respect to the research, authorship, or publication of this article.

References

Aleksandrova, Y., & Parusheva, S. (2019). Social media usage patterns in higher education institutions—An empirical study. International Journal of Emerging Technologies in Learning (iJET), 14(5), 108–121.

Andrews, R., Beynon, M. J., & McDermott, A. M. (2016). Organizational capability in the public sector: A configurational approach. Journal of Public Administration Research and Theory, 26(2), 239–258.

Arora, S., Singh, G. P., Chakraborty, A., & Maity, M. (2022). Polarization and social media: A systematic review and research agenda. Technological Forecasting and Social Change, 183, 121942.

Brunton, G., Thomas, J., O'Mara-Eves, A., Jamal, F., Oliver, S., & Kavanagh, J. (2017). Narratives of community engagement: A systematic review-derived conceptual framework for public health interventions. BMC Public Health, 17(1), 944.

Capriotti, P., Carretón-Ballester, C., & Losada-Díaz, J.-C. (2024). Analysing the influence of Universities' content strategy on the level of engagement on social media. Communication & Society, 37(1), 41–60.

Chen, J. (2024). Impacts of internet literacy and internet contact on the communication effect of university students' ideological and political education in China. Acta Psychologica, 247, 104321.

Chen, S., Li, Q., Lei, B., & Wang, N. (2021). Configurational Analysis of the Driving Paths of Chinese Digital Economy Based on the Technology–Organization–Environment Framework. Sage Open, 11(4), 21582440211054500.

Chu, W. (2023). Research on the innovative path of college students' network ideological and political education. Frontiers in Educational Research, 6(18), 19–24.

Cicchetti, D., & Rogosch, F. A. (1996). Equifinality and multifinality in developmental psychopathology. Development and Psychopathology, 8(4), 597–600.

Dias, M. A., Barreto, M. C. M., Da Silva Cardoso, C., Neto, J. C. L., Pompeu, R. M., & De Alencar, D. B. (2021). University social responsibility and the control mechanisms of the Ministry of Education from a systemic view. International Journal for Innovation Education and Research, 9(1), 22–41.

DiMaggio, P. J., & Powell, W. W. (1983). The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields. American Sociological Review, 48(2), 147–160.

Ding, H. (2022). What kinds of countries have better innovation performance? – A country-level fsQCA and NCA study. Journal of Innovation & Knowledge, 7(4), 100215.

Douglas, E. J., Shepherd, D. A., & Prentice, C. (2020). Using fuzzy-set qualitative comparative analysis for a finer-grained understanding of entrepreneurship. Journal of Business Venturing, 35(1), 105970.

Fiss, P. C. (2011). Building Better Causal Theories: A Fuzzy Set Approach to Typologies in Organization Research. Academy of Management Journal, 54(2), 393–420.

Hinings, B., Gegenhuber, T., & Greenwood, R. (2018). Digital innovation and transformation: An institutional perspective. Information and Organization, 28(1), 52–61.

Isaeva, E. R., Akhmatov, A. A., & Muslieva, T. V. (2024). Innovations in educational technologies as a strategic tool for increasing the competitiveness of educational institutions in the global context. Ekonomika i Upravlenie: Problemy, Resheniya, 2(12), 213–221.

Kandhway, K., & Kuri, J. (2016). Using node centrality and optimal control to maximize information diffusion in social networks. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 47, 1099–1110.

Katz, E., & Lazarsfeld, P. F. (1955). Personal influence: The part played by people in the flow of mass communications. The Free Press.

Kiriliuk, O., & Zakharova, V. (2024). Innovative strategies for the development of the university's ecosystem in the age of digital transformation. Krasnoyarsk Science, 13(2), 239–256.

Kumar, S., Sahoo, S., Lim, W. M., Kraus, S., & Bamel, U. (2022). Fuzzy-set qualitative comparative analysis (fsQCA) in business and management research: A contemporary overview. Technological Forecasting and Social Change, 180, 121599.

Li, W. (2018). Application of virtual reality technology in the inheritance of cultural heritage. Journal of Physics: Conference Series, 1087, 6.

Lian, L. (2023). Research on the Current Situation and Development Strategies of WeChat Official Accounts for Ideological and Political Education in Higher Education Institutions. Media and Communication Research, 4(4), 46–50.

Liu, B., Cui, Z., & Nanyangwe, C. (2023). How line-manager leadership styles and employee-perceived HRM practices contribute to employee performance: A configurational perspective. Leadership & Organization Development Journal, 44(4), 546–563.

Luan, H., Wang, M., Sokol, R., Wu, S., Victor, B., & Perron, B. (2020). A scoping review of WeChat to facilitate professional healthcare education in Mainland China. Medical Education Online, 25(1), 1782594.

Meyer, A. D., Tsui, A. S., & Hinings, C. R. (1993). Configurational approaches to organizational analysis. Academy of Management Journal, 36(6), 1175–1195.

Ngai, C., Singh, R., Lu, W., & Koon, A. (2020). Grappling With the COVID-19 Health Crisis: Content Analysis of Communication Strategies and Their Effects on Public Engagement on Social Media. Journal of Medical Internet Research, 22(8), e21360.

Olya, H. G. T., & Altinay, L. (2016). Asymmetric modeling of intention to purchase tourism weather insurance and loyalty. Journal of Business Research, 69(8), 2791–2800.

Overstreet, M. (2020). Strategies for building community among learners in online courses. College Teaching, 68(1), 45–48.

Pang, H. (2020). Examining associations between university students' mobile social media use, online self-presentation, social support and sense of belonging. Aslib Journal of Information Management, 72(3), 321–338.

Pappas, I. O., Kourouthanassis, P. E., Giannakos, M. N., & Chrissikopoulos, V. (2016). Explaining online shopping behavior with fsQCA: The role of cognitive and affective perceptions. Journal of Business Research, 69(2), 794–803.

Puppatz, M., Wang, M., & Deller, J. (2023). A configurational approach to investigating the relationship between organizational culture and organizational effectiveness using fuzzy-set analysis. Group & Organization Management, 49(4), 1013–1051.

Purtle, J., Marzalik, J. S., Halfond, R. W., Bufka, L. F., Teachman, B. A., & Aarons, G. A. (2020). Toward the data-driven dissemination of findings from psychological science. American Psychologist, 75(8), 1052–1066.

Ragin, C. C. (2006). The calibration of concepts. In B. Rihoux & C. C. Ragin (Eds.), Configurational comparative methods: Qualitative comparative analysis (QCA) and related techniques (pp. 69–86). Sage Publications.

Ragin, C. C. (2008). Redesigning social inquiry: Fuzzy sets and beyond. University of Chicago Press.

Rubin, A. M. (2009). The uses-and-gratifications perspective on media effects. In J. Bryant & M. B. Oliver (Eds.), Media effects: Advances in theory and research (3rd ed., pp. 165–184). Routledge.

Rutter, R., Roper, S., & Lettice, F. (2016). Social media interaction, the university brand and recruitment performance. Journal of Business Research, 69, 3096–3104.

Salmi, J. (2007). Autonomy from the state vs responsiveness to markets. Higher Education Policy, 20(3), 223–242.

Schneider, C. Q., & Wagemann, C. (2012). Set-theoretic methods for the social sciences: A guide to qualitative comparative analysis. Cambridge University Press. https://books.google.com/books?hl=en&lr=&id=sokrhdIi72sC&oi=fnd&pg=PR13&dq=Set-theoretic+methods+for+the+social+sciences:+A+guide+to+qualitative+comparative+analysis&ots=tQpY4wvqqX&sig=gXAMnhn2s8f5oQyqpNB3z0oewu8

Scott, W. R. (2013). Institutions and organizations: Ideas, interests, and identities (4th ed.). Sage Publications.

Shen, L., Wang, S., Chen, W., Fu, Q., Evans, R., Lan, F., Li, W., Xu, J., & Zhang, Z. (2019). Understanding the function constitution and influence factors on communication for the WeChat official account of top tertiary hospitals in China: Cross-sectional study. Journal of Medical Internet Research, 21(9), e13025.

Shu, X. (2023). Research on construction quality evaluation of ideological and political Wechat public accounts in universities under the communication field. SHS Web of Conferences, 171, 03014.

Sun, J., Zhang, J.-H., Zhang, H., Wang, C., Duan, X., & Chen, M. (2020). Development and validation of a tourism fatigue scale. Tourism Management, 81, 104121.

Trueb, B. (2013). Integrating qualitative and quantitative data: Index creation using fuzzy-set QCA. Quality & Quantity, 47(6), 3537–3558.

Valdés-León, K. V., Alvear-Sarrà, S. I., & Suárez-Amaya, L. C. (2021). An institutional perspective for evaluating digital transformation in higher education: Insights from the Chilean case. Sustainability, 13(17), 9850.

Varughese, R. (2017). National policy on education and higher education. Higher Education for the Future, 4(2), 158–165.

Verhoeven, P., & Verbruggen, T. (2024). Digital signaling in public governance: How government agencies use social media to convey legitimacy and performance. Government Information Quarterly, 41(1), 101890.

Weitz, B. (2023). Community engagement. Nature Plants, 9(11), 1777.

Wolfschwenger, J., & Young, K. (2021). Multicausality and equifinality. In K. Young, P. M. F. C. d. B. e. Mello, & G. d. F. C. F. de Mello (Eds.), Research methods in the social sciences: An a-z of key concepts (pp. 209–213). Oxford University Press.

Woodside, A. G. (2014). Embrace• perform• model: Complexity theory, contrarian case analysis, and multiple realities. Journal of Business Research, 67(12), 2495–2503.

Yang, B., Hu, Y., Cheng, X., Bao, Y., & Chen, W. (2023). Exploring the factors affecting content dissemination through WeChat official accounts: A heuristic-systematic model perspective. Electronic Commerce Research, 23(4), 2713–2735.

Yang, S. (2023). A Brief Exploration of the Visual Representation and Communication of Urban Image through Film and Television. Journal of Intelligence and Knowledge Engineering (ISSN: 2959-0620), 1(2), 95.

Yu, Y., & Yao, Y. (2024). Empowerment of precise ideological and political education in higher education with educational digitalization. Journal of Contemporary Educational Research, 8(4), 23–28.

Zhang, H. (2023). A comparative study on discourse and dissemination effect of official WeChat in universities. Media and Communication Research, 4(6), 6–11.

Zhang, M., Wu, D., Chen, H., & Zhou, T. (2024). WeChat official accounts in medical academic libraries: A study on publication status and communication impact in China. The Journal of Academic Librarianship, 50(4), 102946.

Zhang, X., Yan, L., Chen, K., & Sung, D. K. (2020). Fast, efficient broadcast schemes based on the prediction of dynamics in vehicular ad hoc networks. IEEE Transactions on Intelligent Transportation Systems, 21(2), 531–542.

Zheng, X., Cheung, C. M. K., Lee, M. K. O., & Liang, L. (2015). Building brand loyalty through user engagement in online brand communities in social networking sites. Information Technology & People, 28(1), 90–106.

Zhou, Z., Zhang, H., & Sun, Y. (2024). The application and management of university WeChat official accounts in college student ideological education: A exploration based on the “BIFT art sea studio” WeChat official account. SHS Web of Conferences, 187, 01013.