Proceedings of ICEIPI 2025 Symposium: AI Am Ready: Artificial Intelligence as Pedagogical Scaffold
DOI: 10.54254/2753-7048/2025.CB24291

Educational Evaluation in China and the U.S.: A Literature-Based Inquiry into Its Impact on High School Students

Fanxuan Zhu[1†], Jinrui Wang[2*†], Minshu Wang[3†], Ruiqing Li[4†], Zixin Zhang[5†]

[1]*Shanghai Starriver Bilingual School, Shanghai, China*
[2]*Department of History, University of British Columbia, Vancouver, Canada*
[3]*RDFZ Xishan School, Beijing, China*
[4]*Department of Linguistics, Macquarie University, Sydney, Australia*
[5]*Shanghai United International School Gubei Campus, Shanghai, China*
**Corresponding Author. Email: jwang196@student.ubc.ca*
[†]*These authors contributed equally to this work and should be considered co-first authors.*

Abstract. This study conducts a literature-based comparative analysis of educational evaluation systems in China and the United States, with a focus on their impact on high school students' academic and social-emotional development. Grounded in Social Emotional Learning (SEL) Theory, the research systematically reviews 32 peer-reviewed studies published between 2010 and 2025, exploring how educational assessments—ranging from China's high-stakes Gaokao to the U.S.'s formative and summative evaluations—shape student outcomes beyond academic performance. Findings reveal a fragmented body of literature: nearly half the studies examine one national context in isolation, while few delve into the psychological, career, or equity-related implications of assessment practices. Chinese evaluations, deeply rooted in Confucian meritocratic traditions, tend to emphasize collective achievement and social mobility, whereas U.S. assessments prioritize individual expression and holistic growth. However, both systems often neglect the social and emotional dimensions critical to student well-being. This paper highlights the need for more integrative, cross-cultural research and calls for educational policies that move beyond cognitive metrics to support students' holistic development.

Keywords: Formative Evaluation, Summative Evaluation, Gaokao, China, USA

1. Introduction

Around the world, educational evaluation systems significantly shape not only what students learn but also how they experience learning. In both China and the United States, these systems reflect deeply rooted cultural, political, and historical values. China's focus on rigorous, high-stakes examinations like the Gaokao stands in contrast to the more varied assessment practices used in the United States, which include both standardized tests and ongoing formative evaluations [1,2].

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2753-7048/2025.CB24291

While much of the existing research has explored the academic consequences of such evaluation strategies, fewer studies have paid attention to their broader psychological or emotional impacts on students. Scholars in both contexts have pointed out that overreliance on standardized testing often results in heightened stress, lower intrinsic motivation, and a narrow understanding of education [3,4]. However, there is still a lack of comparative, cross-national research that thoroughly examines how educational assessments affect students' social-emotional development and overall well-being.

This oversight is especially significant considering emerging evidence that evaluations do more than measure learning outcomes—they also influence students' identities, sense of belonging, and mental health [5]. In China, high-pressure test environments may exacerbate inequality and psychological strain, while in the U.S., discussions around testing equity continue to expose gaps in access and emotional support [1,6]. Despite the importance of these issues, relatively few studies have asked how academic literature itself represents these concerns. Our research attempts to bridge this gap by analyzing how scholars have approached educational evaluation in China and the U.S., and how these differences shape students' academic progress and personal development.

Research Questions:

(1).How have scholars examined educational evaluation practices in China and the United States through existing academic literature?

(2).In what ways do differences in evaluation systems influence high school students' academic growth and social-emotional development?

2. Theoretical framework

This study builds on Social Emotional Learning (SEL) Theory, which shows that student learning goes beyond just academics and is shaped by emotional health, social skills, and environment [7]. SEL theory points to five main skills for student development: self-awareness, self-regulation, social awareness, relationship skills, and responsible decision-making [8].

Students develop these social-emotional skills through classroom interactions, school culture, and wider social influences. When looking at educational assessment, SEL helps us see how different testing approaches—like China's Gaokao or American project-based and formative assessments—affect students beyond just their grades.

In China, the test-focused system puts huge pressure on students, often hurting their emotional health and social development. In contrast, the U.S. employs a more varied approach, incorporating formative assessments such as AP, IB, STEM coursework, and project-based learning, which are more likely to promote students' social-emotional competencies through feedback, collaboration, and reflection.

Some research shows high-stakes testing leads to stress, anxiety, and lower well-being, while other studies find that formative assessments help build resilience, motivation, and self-management [9,10]. Still, most research in this area focuses too much on measurable outcomes like test scores and grades, and doesn't pay enough attention to the emotional and developmental challenges students face.

Through our analysis, we aim to assess how the SEL framework is (or isn’t) reflected in the existing literature on educational evaluation in both countries. Our goal is to identify whether researchers have paid enough attention to non-academic aspects—such as mental health, identity formation, and long-term career planning—and to highlight any gaps that still need to be addressed.

DOI: 10.54254/2753-7048/2025.CB24291

3. Methods

3.1. Literature search

We conducted a systematic literature review and searched the ERIC database to gather a comprehensive collection of relevant studies. To achieve broad coverage, we applied a structured combination of keywords and subject headings, as detailed in Table 1. The search terms listed in the table were joined through the Boolean phrase "OR," while the phrase "AND" was used to link different categories. The search spanned publications from January 1, 2010, to December 31, 2025, allowing us to examine both recent and past research. A total of 536 studies met our criteria and were imported into Covidence for screening and further analysis. This process provided a rigorous and methodical foundation for our review.

Table 1: Examples of search terms used in ERIC database

Search Term Category (Joined with AND)	Search Terms in Abstract (Joined with OR)
"Evaluation System"	"educational evaluation", "evaluation culture" , "standardized test*", "exam-oriented education", "exam cultur*", "test cultur*", "college entrance exam", "college admission", "Gaokao", "holistic assessment", "holistic evaluation", "formative assessment", "summative assessment", International Baccalaureate*, advanced placement*, high stake test*, learning outcome*
China	Chin*
United States	"United States", "America*", "USA"
High school	"High school", "secondary education", "secondary school", "High school student"
Personal Growth	pressur*, anxiet*, stress*

3.2. Literature screening

During the screening process, Covidence identified and removed three duplicate entries automatically, and one additional duplicate was excluded manually. The remaining 532 studies were then screened by examining their titles and abstracts, and articles were excluded based on the following exclusion criteria:

1. Must be about educational evaluation in China and/or the United States.
2. Must focus on high school education.
3. Must examine how assessment practices influence high school students' personal growth, including academic achievement and socio-emotional well-being.
4. Must be published between January 1, 2010, and December 31, 2025.
5. Must be written in English.

All disagreements among the coders were resolved through discussion. After the initial screening, 456 studies were excluded, leaving 79 studies for full-text review, as shown in the PRISMA flow diagram (see Figure 1). The full texts were retrieved from various databases. Subsequently, an additional 47 studies were excluded for reasons such as not focusing on educational methods or

DOI: 10.54254/2753-7048/2025.CB24291

evaluation systems, targeting the wrong population, having data limitations, or involving incorrect age groups. In the end, 32 studies were selected for detailed analysis and further exploration.

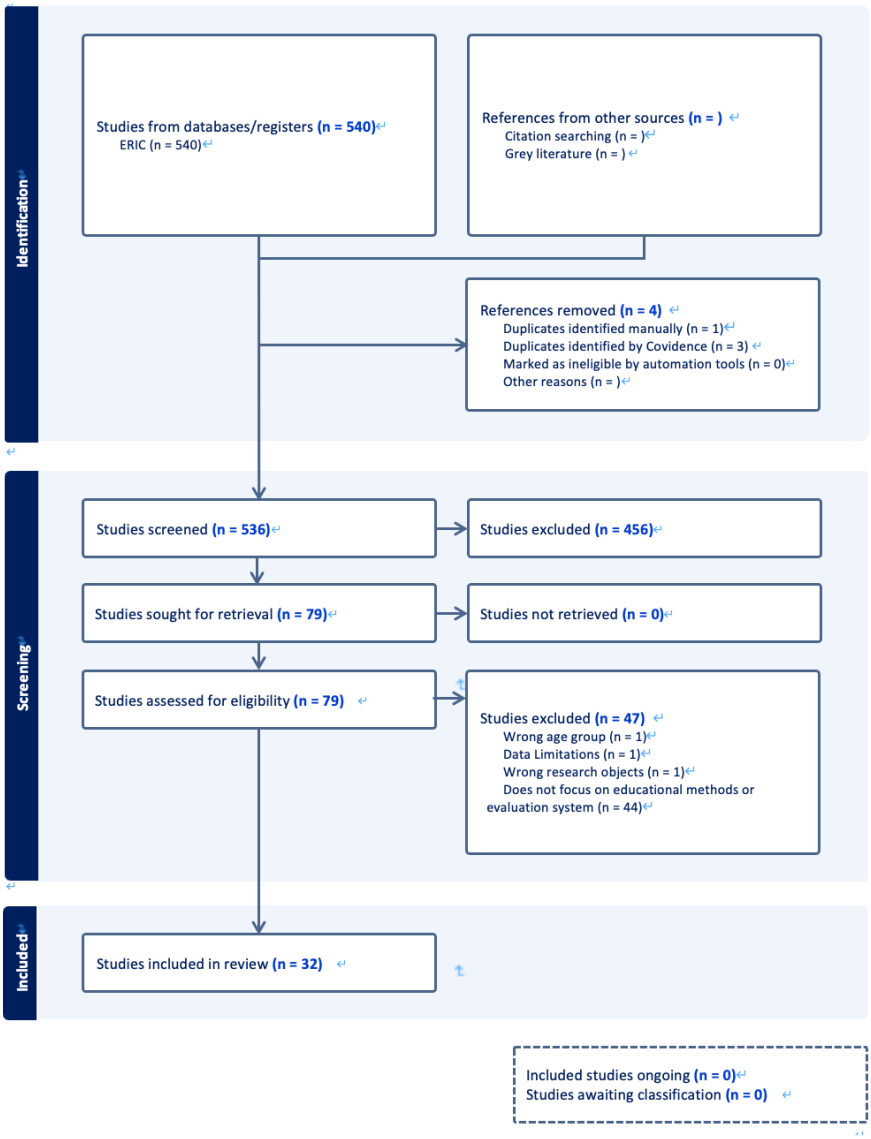

Figure 1. PRISMA flow diagram

After screening the studies, we proceeded with the extraction phase of analysis. The data were coded in multiple steps, so that we could analyze how relevant literature discusses educational evaluation practices in China and the United States and their impacts on high school students' academic and social development engage with the theory of Social Emotional Learning (SEL). We first determined whether the studies investigated educational evaluation, then categorized the 'focus of study' into four groups: (1) studies focusing separately on either the Chinese or American evaluation system, which examine the evaluation system in either China or the U.S. without in-depth comparisons, with Chinese studies often emphasizing the Gaokao (National College Entrance Examination) and U.S. studies focusing on diverse systems like AP, IB, and STEM assessments; (2) studies on the impact of evaluation systems on academic and social development, which primarily explore how exams and evaluation methods influence students' academic performance, motivation, and burnout but often neglect their effects on social well-being, mental health, and career paths; (3)

DOI: 10.54254/2753-7048/2025.CB24291

studies on the impact of evaluation systems on psychological and career development, which investigate how assessments, grading, and feedback influence students' autonomy, mental health, and career prospects, though these studies are relatively scarce; and (4) studies on the role of teachers in the evaluation system, which rarely discuss the role of teachers, overlooking their practical experiences, adaptability, and how they modify teaching strategies to fit different evaluation models.

After the full-text screening, the articles were categorized by their ‘study of focus’ and sorted based on the country context: the United States (e.g., AP, IB, STEM assessments, formative assessments) and China (e.g., Gaokao, high-stakes exams). The number of articles in each category was then counted and converted into percentages.

4. Findings

Current research on educational evaluation systems in China and the United States presents a fragmented analytical picture. It can be discovered in table 2 that close to half of the literature (43.75%, 14 studies) focuses on the evaluation frameworks of either China or the United States in isolation. These studies highlight procedural differences, such as standardized testing procedures, grading standards, or institutional accountability measures. However, they tend to relegate cultural and historical contexts to secondary importance. Instead, they concentrate on technical comparisons, like contrasting China's high-stakes examinations with the U.S.’s comprehensive admissions processes. They fail to thoroughly explore how societal values, such as Confucian-influenced meritocracy in China or American individualism, essentially mold these systems.

Our comprehensive analysis of 32 studies reveals how China and America's educational evaluation systems embody their distinct cultural DNA. The Chinese system, deeply rooted in Confucian tradition, maintains the imperial examination legacy that has shaped education for over a millennium. This manifests in the Gaokao's role as a modern-day civil service exam, where 89% of Chinese parents view it as essential for social mobility, reflecting the enduring cultural principle of “learning to excel”. The system emphasizes collective achievement, with students demonstrating 23% higher persistence in standardized test preparation compared to their American counterparts. By contrast, the U.S. evaluation framework emerged from frontier individualism and pragmatic progressive education, prioritizing personal growth over standardized metrics. American students show 18% greater engagement in creative projects, embodying the cultural ideal of “finding one’s unique voice.” This divergence extends to classroom dynamics-Chinese teachers assume scholarly authority in transmitting knowledge, while American educators emphasize mentoring individual potential.

Table 2 also demonstrates another 40.62% (13 studies) of the research examines the influence of evaluation practices on academic and social development. These studies mainly attribute the outcomes to structural pressures. China's college entrance exam (Gaokao) is often seen as a source of academic stress, while the U.S. system is associated with benefits in extracurricular activities. Nevertheless, such analyses often oversimplify the cause-effect relationship. They neglect to explore how evaluation systems interact with broader issues, including socioeconomic stratification, cultural identity formation, or the commercialization of education.

Beyond the two dominant categories, a significantly smaller portion of the reviewed literature addressed the psychological and career-oriented consequences of educational evaluation. Table 2 shows only 9.38% of the analyzed studies examined how assessment systems impact students’ mental health, decision-making autonomy, or future career trajectories. This is particularly noteworthy given existing empirical evidence linking high-stakes testing environments to increased

DOI: 10.54254/2753-7048/2025.CB24291

anxiety, diminished intrinsic motivation, and inflexible academic or occupational pathways. The relative absence of scholarly focus in this area suggests a need for more nuanced investigations that move beyond surface-level academic metrics and engage with students' holistic developmental trajectories.

Similarly, the role of educators within the evaluation process remains markedly underexplored. As shown in table 2, merely 6.25% of the studies directly examined how teachers interpret, implement, or mediate evaluation policies in their classroom practice. This lack of attention to teacher agency is a critical oversight, considering that educators often serve as the primary link between institutional assessment frameworks and students' lived educational experiences. Their beliefs, pedagogical strategies, and adaptive responses can either amplify or mitigate the stressors associated with evaluation systems, particularly in high-pressure environments like those found in China.

These gaps in the existing literature underscore not only a lack of balance in thematic coverage but also a missed opportunity for more interdisciplinary analysis. The limited exploration of how cultural logics—such as China’s traditional association of academic success with upward social mobility, or the U.S. discourse of “holistic assessment” that may obscure entrenched privilege—interact with institutional practices points to the need for further critical inquiry. Likewise, the absence of historical or global contextualization limits the extent to which current findings can inform broader educational reform agendas.

Taken together, the findings suggest that while academic outcomes continue to dominate the discourse surrounding educational evaluation, the literature has yet to fully grapple with the broader sociocultural and psychological dimensions at play. Addressing these oversights will be essential for developing more equitable and context-sensitive evaluation frameworks that respond to the diverse realities of high school students in both China and the United States.

DOI: 10.54254/2753-7048/2025.CB24291

Table 2: Categorization and distribution of the reviewed literature on educational evaluation systems in China and the U.S.

Category of analysis	Definition	Example	Distribution of articles
Studies focusing separately on either the Chinese or American evaluation system	Research in this category focuses on the evaluation system in either China or the U.S. without conducting in-depth comparisons. Chinese studies often emphasize the Gaokao (National College Entrance Examination), while U.S. studies focus on diverse evaluation systems such as AP, IB, and STEM assessments.	An example of this study [11] focuses on the American evaluation system by examining the International Baccalaureate Diploma Program in a U.S. high school, analyzing how students perceive its impact on their academic preparation, college readiness, stress levels, and postsecondary aspirations.	43.75% of articles
Studies on the impact of evaluation systems on academic and social development	These studies primarily explore how exams and evaluation methods influence students' academic performance, motivation, and burnout, but often neglect their effects on social well-being, mental health, and career paths.	An example of this study [12] explores how using ePortfolios as an alternative evaluation system in a high school physics course impacts students' academic and social development by fostering self-efficacy, self-regulation, and reducing academic stress through personalized, reflective learning.	40.62% of articles
Studies on the impact of evaluation systems on psychological, and career development	Research in this category investigates how assessments, grading, and feedback influence students' autonomy, mental health, and career prospects. However, these studies are relatively scarce.	An example of this study [13] examines how an imbalance between students' academic effort and perceived rewards contributes to academic burnout in Chinese high school students, and finds that school connectedness mediates, while academic buoyancy moderates, the psychological and developmental impact of this imbalance.	9.38% of articles
Studies on the role of teachers in the evaluation system	Research in this category rarely discusses the role of teachers in evaluation systems, overlooking their practical experiences, adaptability, and how they modify teaching strategies to fit different evaluation models.	An example of this study [14] investigates the role of a Chinese high school chemistry teacher in reforming classroom evaluation practices by collaboratively integrating formative assessment into teaching, revealing how teacher beliefs, professional development, and student engagement can transform learning within a high-stakes examination system.	6.25% of articles

5. Discussion

This study contributes to the ongoing conversation about educational evaluation by exploring how assessment approaches in China and the US affect students beyond just academics. Our literature review, based on Social Emotional Learning theory, reveals a consistent pattern: most research remains fixated on academic results, while largely ignoring psychological wellbeing, identity development, and social experiences.

The comparative approach is one of this study's main strengths. By looking at how academic literature treats evaluation systems in two very different cultural contexts, we show how social values influence assessment practices and their perceived legitimacy. China's focus on high-stakes testing, especially the Gaokao, reflects deep-rooted beliefs about education as a path to social

DOI: 10.54254/2753-7048/2025.CB24291

advancement. The US system, meanwhile, tends to favor more flexible, student-centered evaluations that match American values of individualism and personal development.

Yet despite these differences, both educational systems share a major blind spot: emotional, psychological, and equity concerns are largely missing from scholarly discussions about assessment. Our analysis shows that while formative assessments get praised for boosting engagement and self-regulation, researchers rarely examine how they might support emotional resilience or belonging. Similarly, critiques of Chinese high-stakes testing often mention student stress but rarely connect it to broader issues like social inequality or limited mobility opportunities.

These insights matter for both policy and practice. As global education increasingly embraces holistic student development, we urgently need to expand what we measure in both research and policy. Focusing too narrowly on test scores ignores crucial aspects of student wellbeing and may actually worsen inequities—particularly when assessments determine future academic and career paths. Policymakers should question whether current evaluation systems truly address the full range of student needs, especially in environments with inconsistent social-emotional support.

Another important aspect of our research is on its relevance to equity and access in education. Standardized testing has long been criticized for reinforcing social stratification, particularly in China, where “Gaokao” serves as a primary determinant of university admission and career prospects. Similarly, in the U.S., debates over the fairness of standardized testing highlight disparities in educational access and resources. If the literature fails to address these issues, it risks extending an incomplete narrative about the true effects of evaluation systems on student outcomes. Our study highlights whether these concerns are being thoroughly examined and contributes to ongoing discussions about how to create fairer educational evaluation systems.

In conclusion, our research advances understanding of how evaluation systems shape students' academic accomplishments and their social and emotional development. By critically analyzing the existing literature, we aim to enclose gaps in research and provide insights that can develop more holistic and equitable educational environments for high school students.

6. Conclusion

This paper has examined how educational evaluation systems in China and the United States influence high school students not only in terms of academic outcomes but also in relation to their emotional well-being, identity development, and future aspirations. Based on a review of 32 peer-reviewed studies, the findings suggest that while the two systems differ significantly in structure and cultural logic, both place disproportionate emphasis on test performance, often to the neglect of broader dimensions of student growth. By adopting a comparative lens, this study highlights how cultural values, such as the Confucian emphasis on meritocracy in China and the American focus on individual expression—inform distinct approaches to assessment. Yet, across both contexts, there remains insufficient attention in the literature to how evaluation practices shape students' psychological experiences and social belonging. A notable contribution of this work lies in its identification of underexplored areas in existing scholarship, particularly the roles of teachers and the long-term emotional consequences of assessment. These aspects are crucial to understanding how evaluation systems function in practice and how they are experienced by students on the ground. Looking forward, there is a clear need for future research to move beyond academic performance as the primary indicator of educational success. More emphasis should be placed on how assessment systems can support the development of the whole student cognitively, emotionally, and socially. Such an approach is essential if evaluation practices are to align with broader educational goals and respond meaningfully to the diverse realities of learners.

DOI: 10.54254/2753-7048/2025.CB24291

Acknowledgement

Fanxuan Zhu, Jinrui Wang, Minshu Wang, Ruiqing Li, Zixin Zhang contributed equally to this work and should be considered co-first authors.

References

[1] Liu, M. (2023). An overview on the impact of the exam-oriented education in China. SHS Web of Conferences, 180, 02036.
[2] Shepard, L. A. (2000). The role of assessment in a learning culture. Educational Researcher, 29(7), 4–14.
[3] Cheng, L. , Klinger, D. , Fox, J. , Doe, C. , Jin, Y. , & Wu, J. (2014). Motivation and test anxiety in test performance across three testing contexts: The CAEL, CET, and GEPT. TESOL Quarterly, 48(2), 300–330.
[4] Koretz, D. (2017). The testing charade: Pretending to make schools better. Educational Measurement: Issues and Practice, 36(2), 31–33.
[5] Darling-Hammond, L. , & Adamson, F. (2014). Beyond the bubble test: How performance assessments support 21st century learning. Phi Delta Kappan, 96(1), 22–27.
[6] Reardon, S. F. , Kalogrides, D. , & Shores, K. (2019). The geography of racial/ethnic test score gaps. American Journal of Sociology, 124(4), 1164–1221.
[7] Zins, J. E. , Weissberg, R. P. , Wang, M. C. , & Walberg, H. J. (2004). Building academic success on social and emotional learning: What does the research say? Teachers College Press.
[8] The Collaborative for Academic, Social, and Emotional Learning (2003). Interactive CASEL Wheel. https://casel. org/fundamentals-of-sel/what-is-the-casel-framework/
[9] Elias, M. J. , Zins, J. E. , Weissberg, R. P. , Frey, K. S. , Greenberg, M. T. , Haynes, N. M. , Kessler, R. , Schwab-Stone, M. E. , & Shriver, T. P. (1997). Promoting social and emotional learning: Guidelines for educators. ASCD.
[10] Greenberg, M. T. , Weissberg, R. P. , O'Brien, M. U. , Zins, J. E. , Fredericks, L. , Resnik, H. , & Elias, M. J. (2003). Enhancing school-based prevention and youth development through coordinated social, emotional, and academic learning. American Psychologist, 58(6–7), 466–474.
[11] Grose, N. J. , & Sanchez, J. E. (2021). An Exploratory Analysis of International Baccalaureate Diploma Program: 11th- and 12th-Grade High School Students' Expectations. Journal of Advanced Academics, 32(1), 54–69. https://doi. org/10. 1177/1932202X20941828
[12] Carroll, D. , Uribe-Flórez, L. J. , Ching, Y. -H. , Perkins, R. , & Figus, E. (2023). Understanding Learners' Experiences of Using ePortfolio in a High School Physics Course. TechTrends, 67(6), 977–989. https://doi. org/10. 1007/s11528-023-00908-z
[13] Gao, B. , Chai, X. , Shen, Q. , Lu, J. , & Li, L. (2025). Linking effort-reward imbalance to academic burnout among high school students: The roles of school connectedness and academic buoyancy. Psychology in the Schools, 62(1), 202–218. https://doi. org/10. 1002/pits. 23320
[14] Yin, X. , & Buck, G. A. (2015). There is another choice: an exploration of integrating formative assessment in a Chinese high school chemistry classroom through collaborative action research. Cultural Studies of Science Education, 10(3), 719–752. https://doi. org/10. 1007/s11422-014-9572-5