

SYSTEMATIC REVIEW

## Assessment and Evaluation in the Senior High School: A Systematic Literature Network Analysis

### Evaluación y valoración en la escuela secundaria superior: un análisis sistemático de la red bibliográfica

Mukhamad Zulianto<sup>1,2</sup> , Arief Budiono<sup>3,4</sup>  , Irmadatus Sholekhah<sup>2</sup> , Arum Putri Rahayu<sup>5</sup> , Turdialiev Mukhammad Ali Polatjon ogli<sup>6</sup> 

<sup>1</sup>National Dong Hwa University, Asia-Pacific Regional Studies. Hualien, Taiwan, R.O.C.

<sup>2</sup>Universitas Jember, Department of Economics Education. Jember, Indonesia.

<sup>3</sup>Universitas Muhammadiyah Surakarta, Department of Doctoral Law. Surakarta, Indonesia.

<sup>4</sup>Tashkent State University of Law, Guest Professor. Tashkent, Uzbekistan.

<sup>5</sup>STAI Ma'arif Magetan, Department of Islamic Education. Magetan, Indonesia.

<sup>6</sup>Tashkent State University of Law, Department of Private International Law. Tashkent, Uzbekistan.

**Cite as:** Zulianto M, Budiono A, Sholekhah I, Putri Rahayu A, Ali Polatjon ogli TM. Assessment and Evaluation in the Senior High School: A Systematic Literature Network Analysis. Data and Metadata. 2025; 4:1250. <https://doi.org/10.56294/dm20251250>

Submitted: 28-06-2025

Revised: 22-09-2025

Accepted: 04-11-2025

Published: 05-11-2025

Editor: Dr. Adrián Alejandro Vitón Castillo 

Corresponding author: Arief Budiono 

#### ABSTRACT

**Introduction:** assessment and evaluation have become central to shaping educational quality in senior high schools, reflecting global reforms that emphasize competency-based frameworks and holistic learning.

**Method:** the study uses a Systematic Literature Network Analysis (SLNA), combining bibliometric analysis with a Systematic Literature Review (SLR). This study aimed to describe the dominant themes, methodological patterns, and innovations in assessment practices between 2015 and 2025, while identifying persistent challenges and opportunities for reform. The 43 peer-reviewed articles were selected from the Scopus database based on strict inclusion criteria, with the data thematically analysed and visualised using bibliometric network mapping.

**Results:** the findings reveal a marked shift from traditional, summative models toward competency-based assessment, with increasing attention to higher-order thinking skills, reflective strategies, and student self-assessment. Teacher agency and professional development emerged as critical factors in shaping assessment cultures, although systemic challenges such as policy-practice misalignment, limited training, and the digital divide constrain implementation. This study contributes to the existing body of knowledge by synthesizing pedagogical, psychological, and organizational perspectives, thereby framing assessment as a multidimensional construct that drives educational equity and quality. The results underscore the importance of culturally responsive and contextually adaptable practices.

**Conclusions:** future research should explore scalability, sustainability, and cross-context validation to ensure that assessment reforms in senior high schools remain effective and equitable.

**Keywords:** Assessment; Evaluation; Senior High School; Systematic Literature Review.

#### RESUMEN

**Introducción:** la evaluación se ha convertido en un elemento central para configurar la calidad educativa en los institutos de secundaria, lo que refleja las reformas globales que hacen hincapié en los marcos basados en competencias y el aprendizaje holístico.

**Método:** el estudio utiliza un análisis sistemático de redes bibliográficas (SLNA), que combina el análisis bibliométrico con una revisión sistemática de la literatura (SLR). El objetivo de este estudio era describir los temas dominantes, los patrones metodológicos y las innovaciones en las prácticas de evaluación entre 2015 y 2025, al tiempo que se identificaban los retos persistentes y las oportunidades de reforma. Se seleccionaron 43 artículos revisados por pares de la base de datos Scopus basándose en criterios de inclusión estrictos, y los datos se analizaron temáticamente y se visualizaron mediante mapas bibliométricos de red.

**Resultados:** los resultados revelan un marcado cambio de los modelos tradicionales y sumativos hacia la evaluación basada en competencias, con una atención creciente a las habilidades de pensamiento de orden superior, las estrategias reflexivas y la autoevaluación de los estudiantes. La agencia docente y el desarrollo profesional surgieron como factores críticos en la configuración de las culturas de evaluación, aunque los retos sistémicos, como la desalineación entre las políticas y la práctica, la formación limitada y la brecha digital, limitan su implementación. Este estudio contribuye al conjunto de conocimientos existentes al sintetizar las perspectivas pedagógicas, psicológicas y organizativas, enmarcando así la evaluación como un constructo multidimensional que impulsa la equidad y la calidad educativas.

**Conclusiones:** las investigaciones futuras deberían explorar la escalabilidad, la sostenibilidad y la validación entre contextos para garantizar que las reformas de la evaluación en los institutos de secundaria sigan siendo eficaces y equitativas.

**Palabras clave:** Valoración; Evaluación; Bachillerato; Revisión Sistemática de la Literatura.

## INTRODUCTION

The value of assessment and evaluation has emerged as a key ingredient in shaping the quality of educational experiences, particularly in senior high school institutions. As worldwide academic structures transform to highlight proficiency-driven curricula and enduring learning ideologies.<sup>(1)</sup> The methodologies used for assessing students' progress and the effectiveness of instructional practices are currently undergoing a critical reassessment.<sup>(2)</sup> The traditional frameworks that emphasized summative evaluation are presently experiencing a shift or enhancement through the incorporation of formative, authentic, and reflective assessment approaches, which are explicitly designed to advance not only academic success but also to cultivate critical thinking, creativity, and transferable skills.<sup>(3)</sup> Poverty eradication, zero hunger, good health and well-being, quality education, gender equality, clean water and sanitation, and many other issues that are now the focus of the Sustainable Development Goal (SDGs).<sup>(4,5)</sup> This profound transformation is distinctly observable when scrutinized within the context of SDGs 4, which highlights the pivotal significance of securing that inclusive and high-quality education is reachable for all individuals, consequently enhancing equitable pathways for learning and individual progress among varied populations and communities.<sup>(6)</sup>

A thorough investigation of the scholarly works shows a variety of research efforts dedicated to pioneering evaluation strategies in secondary schooling. A study also discusses the application of Data Envelopment Analysis (DEA) as a methodological tool for assessing efficiency in resource allocation in high schools, adopting evaluation as an approach to organizational analysis.<sup>(7)</sup> Research in science education reveals that authentic assessment models such as the Statistical Analysis System (SAS) have been validated for their effectiveness and practicality in assessing student competence.<sup>(8)</sup> Examined the self-evaluation methodologies employed by educators within the framework of the ADDIE model for biology education, highlighting the significance of reflective professionalism.<sup>(9)</sup> Analyzed teachers' attitudes toward Indonesia's national competency-based assessment (Asesmen Kompetensi Minimum - AKM).<sup>(10)</sup> Emphasizing structural opposition arising from inadequate preparedness, these analyses illuminate the augmented complexity and importance of evaluation as a multi-dimensional occurrence that has pedagogical, psychological, and organizational ramifications.<sup>(11)</sup>

The expanding corpus of research, significant obstacles remain concerning the execution and conceptual precision of assessment and evaluation within the senior high school context.<sup>(1)</sup> A longstanding issue relates to the incongruity between the frameworks instituted at the policy level and the actual experiences encountered within educational environments.<sup>(12)</sup> National assessment reforms may promote the development of higher-order thinking skills (HOTS), the prevailing pedagogical practices within classrooms frequently continue to be entrenched in memorization and basic cognitive questioning.<sup>(13)</sup> The capability of educators to adapt and implement innovative evaluative strategies remains variable, particularly within educational systems that are under-resourced or decentralized.<sup>(14)</sup> The imbalance in digital availability and the failings of development frameworks amplify the struggles connected with these implementations.<sup>(15)</sup> The first section introduces the selection process for publications on Assessment and Evaluation in Senior High Schools and the results obtained from bibliometric analysis, which identifies five main clusters in the literature.

## METHOD

### Research Design

This study uses the Systematic Literature Network Analysis (SLNA) methodology, which combines qualitative and quantitative approaches, providing in-depth information on the topics discussed.<sup>(16)</sup> The SLNA approach is suitable for synthesizing fragmented literature and offering a comprehensive understanding of complex topics, such as assessment and evaluation practices in senior high schools. It emphasizes methodological transparency, replicability, and structured synthesis of existing empirical evidence. This study uses PRISMA as an approach, a systematic review of the literature, and VOSviewer software. Which ensures methodological rigor through clearly defined inclusion criteria, systematic data collection, and structured reporting.

### Data Source and Literature Search

The sources identified as critical and consequential to the thorough systematic review we are presently conducting have been solely obtained from the Scopus database. It is regarded and appreciated as one of the most extensive and frequently consulted bibliographic resources in the academic and scholarly realm, thus furnishing a reliable basis for the research findings we seek to disseminate. Scopus offers the most consistent, exportable citation metadata and broad education coverage, yielding cleaner, more reproducible networks. Scopus provides the most consistent, exportable citation metadata and broad education coverage, yielding cleaner, more reproducible networks. Scopus was chosen for its comprehensive multidisciplinary coverage and its incorporation of peer-reviewed scholarly articles. The investigation of the literature was executed in August 2025, employing various combinations of the subsequent keywords: “assessment,” “evaluation,” and “senior high school.” This study uses search strings in “Title” and “Keywords” to obtain: (TITLE-ABS-KEY(assessment) AND TITLE-ABS-KEY(evaluation) AND TITLE-ABS-KEY(senior high school)) AND ( LIMIT-TO ( LANGUAGE, “English” ) ) AND ( LIMIT-TO ( DOCTYPE, “ar” ) ) AND ( LIMIT-TO ( SRCTYPE, “j” ) ) AND PUBYEAR > 2015 AND PUBYEAR < 2025.

### Inclusion and Exclusion Criteria

The researchers established criteria for article selection: (a) articles must be scientific articles that have undergone peer review, excluding literature findings, book chapters, conference proceedings, research reports, books, dissertations, and theses, (b) articles must be indexed in the Scopus database, with a publication date range from 2015 to 2025, (c) articles must be written in English, (d) the subject area of articles is limited to social sciences, multidisciplinary, arts and humanities, and psychology, and (f) articles must be published in open access journals.

## RESULTS

The article search process utilizes the PRISMA flowchart, employing four stages: identification, screening, eligibility, and inclusion. These stages are illustrated in figure 1 below.

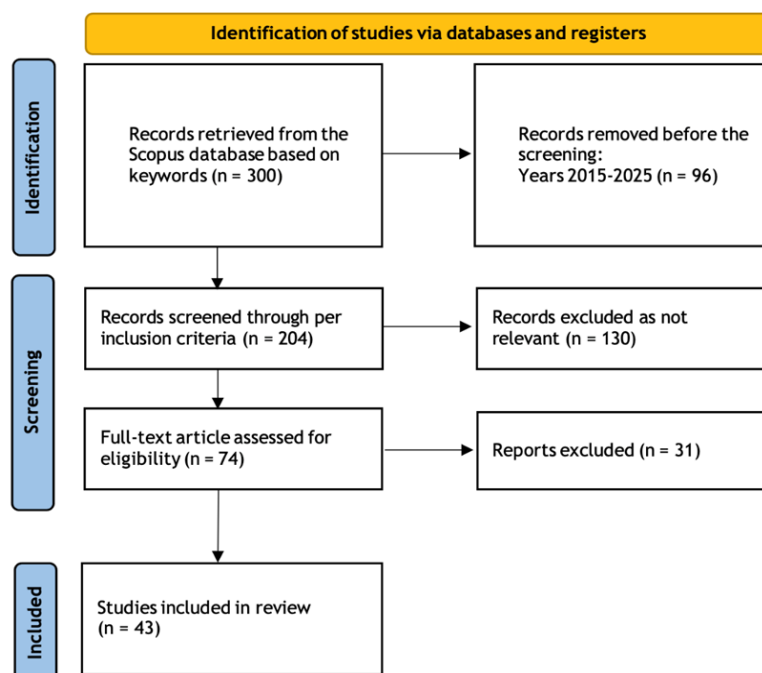


Figure 1. The PRISMA Flow Diagram

**Table 1.** The findings of 43 articles from the Scopus databases

| No. | Journals                                                      | Authors | Year |
|-----|---------------------------------------------------------------|---------|------|
| 1   | PLOS One                                                      | (17)    | 2025 |
| 2   | Educational Process International Journal                     | (18)    | 2025 |
| 3   | Jurnal Ilmiah Peuradeun                                       | (12)    | 2025 |
| 4   | Heliyon                                                       | (19)    | 2024 |
| 5   | International Journal of Evaluation and Research in Education | (9)     | 2024 |
| 6   | Biosaintifika                                                 | (11)    | 2024 |
| 7   | Journal of Surgical Education                                 | (20)    | 2024 |
| 8   | Cakrawala Pendidikan                                          | (10)    | 2024 |
| 9   | Retos                                                         | (21)    | 2024 |
| 10  | Journal of Education and Learning                             | (22)    | 2023 |
| 11  | IJERE                                                         | (23)    | 2023 |
| 12  | International Journal of Instruction                          | (24)    | 2023 |
| 13  | International Journal of Education and Practice               | (8)     | 2023 |
| 14  | Frontiers in Education                                        | (25)    | 2023 |
| 15  | Journal of Educational and Social Research                    | (26)    | 2023 |
| 16  | Journal of Technology and Science Education                   | (2)     | 2023 |
| 17  | Journal of Education and Health Promotion                     | (14)    | 2023 |
| 18  | Advanced Education                                            | (15)    | 2022 |
| 19  | International Journal of Evaluation and Research in Education | (27)    | 2022 |
| 20  | Thinking Skills and Creativity                                | (28)    | 2022 |
| 21  | Jurnal Pendidikan Agama Islam                                 | (7)     | 2022 |
| 22  | Journal of Curriculum and Teaching                            | (29)    | 2022 |
| 23  | Perspectives of Science and Education                         | (30)    | 2022 |
| 24  | BMC Medical Education                                         | (6)     | 2021 |
| 25  | PLOS One                                                      | (31)    | 2021 |
| 26  | Research on Social Work Practice                              | (32)    | 2021 |
| 27  | European Journal of Educational Research                      | (3)     | 2021 |
| 28  | PNAS                                                          | (33)    | 2020 |
| 29  | Cogent Education                                              | (34)    | 2020 |
| 30  | Codas                                                         | (25)    | 2020 |
| 31  | Journal for the Education of Gifted Young Scientists          | (35)    | 2020 |
| 32  | Iafor Journal of Education                                    | (1)     | 2020 |
| 33  | PLOS One                                                      | (36)    | 2020 |
| 34  | Journal on Mathematics Education                              | (37)    | 2019 |
| 35  | Jurnal Pendidikan Islam                                       | (38)    | 2019 |
| 36  | Journal for the Education of Gifted Young Scientists          | (39)    | 2019 |
| 37  | International Journal of Instruction                          | (40)    | 2018 |
| 38  | Universal Journal of Educational Research                     | (13)    | 2018 |
| 39  | BMC Medical Education                                         | (41)    | 2018 |
| 40  | Journal on Mathematics Education                              | (42)    | 2018 |
| 41  | BMC Medical Education                                         | (43)    | 2018 |
| 42  | PLOS One                                                      | (44)    | 2018 |
| 43  | Psychology in Russia State of the Art                         | (45)    | 2015 |

The identification stage, a total of 300 articles indexed in Scopus were initially found. Furthermore, at the screening stage, the articles underwent a publication focus only assessment in 2015-2025, with 204 different articles retained. In the next screening, 130 irrelevant articles were excluded, leaving 74 articles for further consideration. After reaching the eligibility stage, 74 articles were selected for comprehensive reading, while 31 articles were discarded. Finally, 43 articles were selected, with the subject area of articles limited to social sciences, multidisciplinary, arts and humanities, and psychology from 74 eligible articles, aligned with the

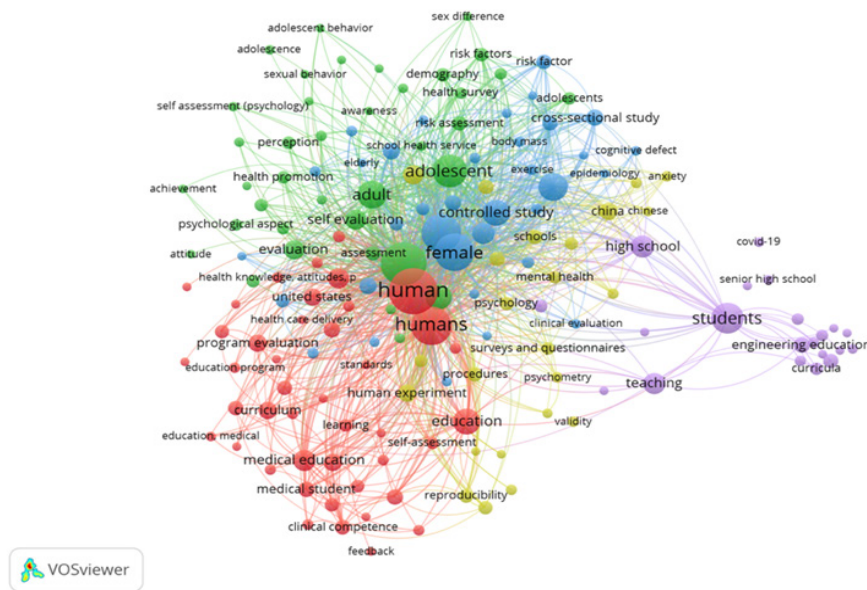
research questions and met the relevance criteria in terms of title, abstract, keywords, and content, while the remaining 31 articles were not used.

### Network visualization

The bibliometric visualization generated using VOSviewer reveals a structured network of keyword co-occurrences within the body of literature related to “Assessment and Evaluation in Senior High School.” Through a systematic literature network analysis, several distinct clusters emerge, each representing a thematic concentration of research. These clusters provide insight into the scope, orientation, and interconnections of scholarly work in this domain. The analysis identified five major clusters, each color-coded and spatially grouped based on shared keywords and semantic proximity. These clusters (figure 2) represent disciplinary foci ranging from medical education and psychometrics to adolescent mental health and high school pedagogy. The circles network visualization results of the VOSViewer software in figure 2 show that there are five clusters related, including:

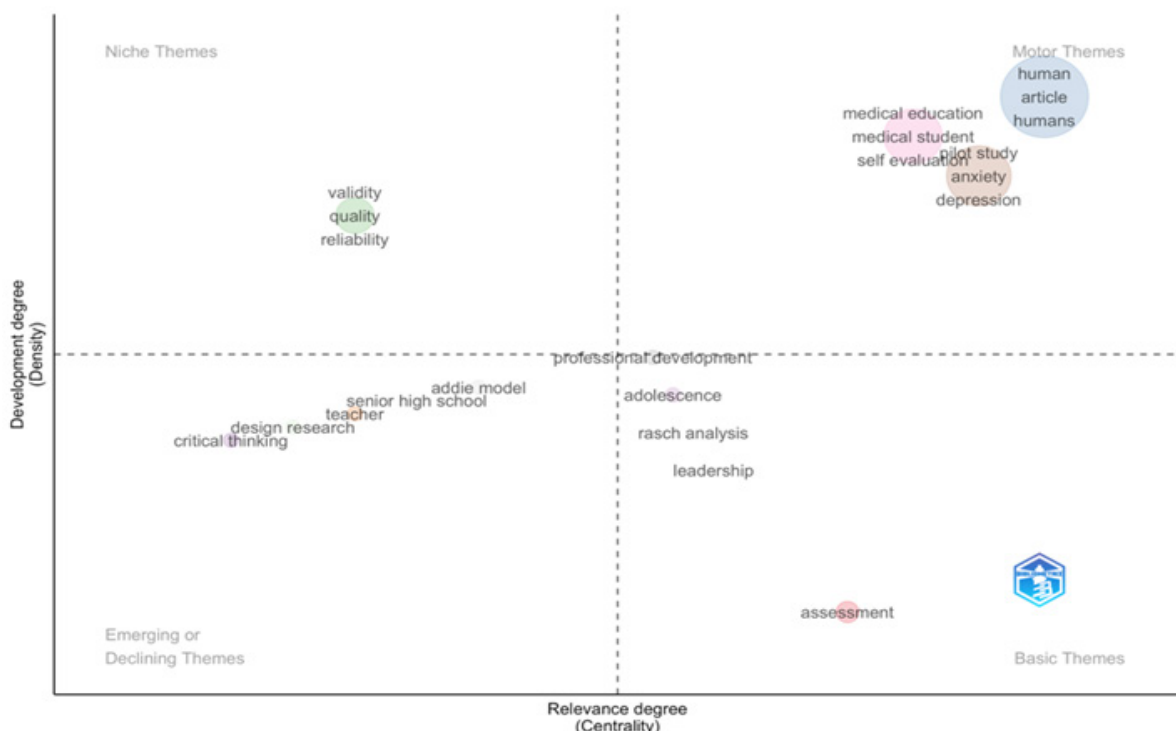
- Cluster 1 (Red): Medical and Programmatic Evaluation in Health Education: the red cluster, anchored by the central terms “human,” “program evaluation,” and “medical education,” encapsulates a large body of literature focused on evaluating medical education and healthcare-related learning programs. Prominent keywords include “clinical competence,” “feedback,” “medical student,” “curriculum,” “education program,” and “healthcare delivery.”
- Cluster 2 (blue): Adolescent Health and School-Based Evaluation Studies: the blue cluster is primarily centered on the adolescent demographic and school-based empirical studies. Key terms in this cluster include “adolescent,” “female,” “controlled study,” “cross-sectional study,” “mental health,” “exercise,” and “schools.” This group of studies applies health and psychological evaluations within secondary school settings, reflecting an intersection of education, public health, and social science research. Many studies in this cluster investigate the psychological well-being, risk behaviors, and lifestyle factors of adolescents—issues highly relevant to senior high school environments. The prevalence of methodologies such as controlled trials and cross-sectional designs suggests a strong orientation toward empirical validation of interventions and program impacts. The inclusion of keywords like “anxiety,” “cognitive defect,” and “risk assessment” indicates a growing research interest in how mental health affects learning.
- Cluster 3 (Green): Psychological and Self-Assessment Constructs: the green cluster focuses on psychological constructs and self-evaluative mechanisms. Keywords such as “self-assessment,” “self-evaluation,” “perception,” “awareness,” “attitude,” and “adolescent behavior” dominate this cluster. These expressions characterize a theoretical model that is anchored in metacognitive understanding, behavioral approaches, and the identity growth processes, particularly for learners in their teenage years. The methodologies of self-assessment and reflective learning are increasingly recognized as integral elements of contemporary pedagogical approaches, particularly within paradigms that promote learner autonomy and iterative feedback systems. In the context of senior high schools, this grouping highlights the critical necessity of nurturing students’ self-regulation and evaluative judgment capabilities as integral elements of holistic educational assessments.
- Cluster 4 (Yellow): Psychometrics, Instrumentation, and Methodological Validity: the yellow cluster is characterized by a methodological orientation, encompassing pivotal terminologies such as “validity,” “psychometry,” “surveys and questionnaires,” “procedures,” and “reproducibility.” This cluster encompasses scholarly literature concentrated on the development, validation, and reliability of evaluative instruments. The availability of a psychometric lexicon denotes a methodical, evidence-supported framework for the formulation and appraisal of tools used in assessing educational results, psychological features, and an array of academic constructs. For educators and researchers at the senior high school level, a thorough comprehension of the statistical and theoretical foundations of validity and reliability is imperative when utilizing rubrics, standardized assessments, or diagnostic tools.
- Cluster 5 (Purple): Senior High School, Curriculum, and Student-Centered Evaluation: the purple cluster represents the thematic core of this study, explicitly focused on high school education. It contains terms such as “students,” “high school,” “senior high school,” “teaching,” “engineering education,” “curricula,” and “COVID-19.” Unlike the other clusters, which may inform high school assessment indirectly, this cluster directly engages with the institutional, curricular, and pedagogical dimensions of senior high school education. Then “engineering education” reflects a growing emphasis on STEM (Science, Technology, Engineering, and Mathematics) in secondary education. The appearance of “COVID-19” as a keyword further signals the shift in teaching and evaluation practices brought about by the pandemic—particularly the challenges and innovations in remote assessment. Within this cluster, the focus is clearly on curriculum implementation, teaching strategies, and student performance both in academic and applied learning contexts. This cluster forms the empirical and conceptual backbone for exploring real-world assessment practices in senior high schools.





**Figure 2.** Network visualization

The thematic map provides a visual representation of the distribution of research themes identified through a systematic literature network analysis. It is structured along two axes: the horizontal axis represents the relevance degree (centrality), indicating how important a theme is within the overall network, while the vertical axis shows the development degree (density), reflecting how well-developed or mature a theme is in existing literature. Themes in the upper-right quadrant, known as motor themes, are both central and well-developed, including human, medical education, self-evaluation, and medical student, which demonstrate strong research attention and interconnectivity.<sup>(46,47)</sup> The upper-left quadrant contains niche themes like validity, reliability, and quality, representing well-developed but more isolated areas of study. In the lower-left quadrant are emerging or declining themes, such as critical thinking and design research, which are neither central nor well-developed, possibly indicating outdated or still-developing topics. Finally, the lower-right quadrant presents basic themes, such as assessment, which are relevant but still underdeveloped. This map offers strategic insight into dominant trends, research gaps, and future directions in the field of senior high school assessment and evaluation (figure 3).

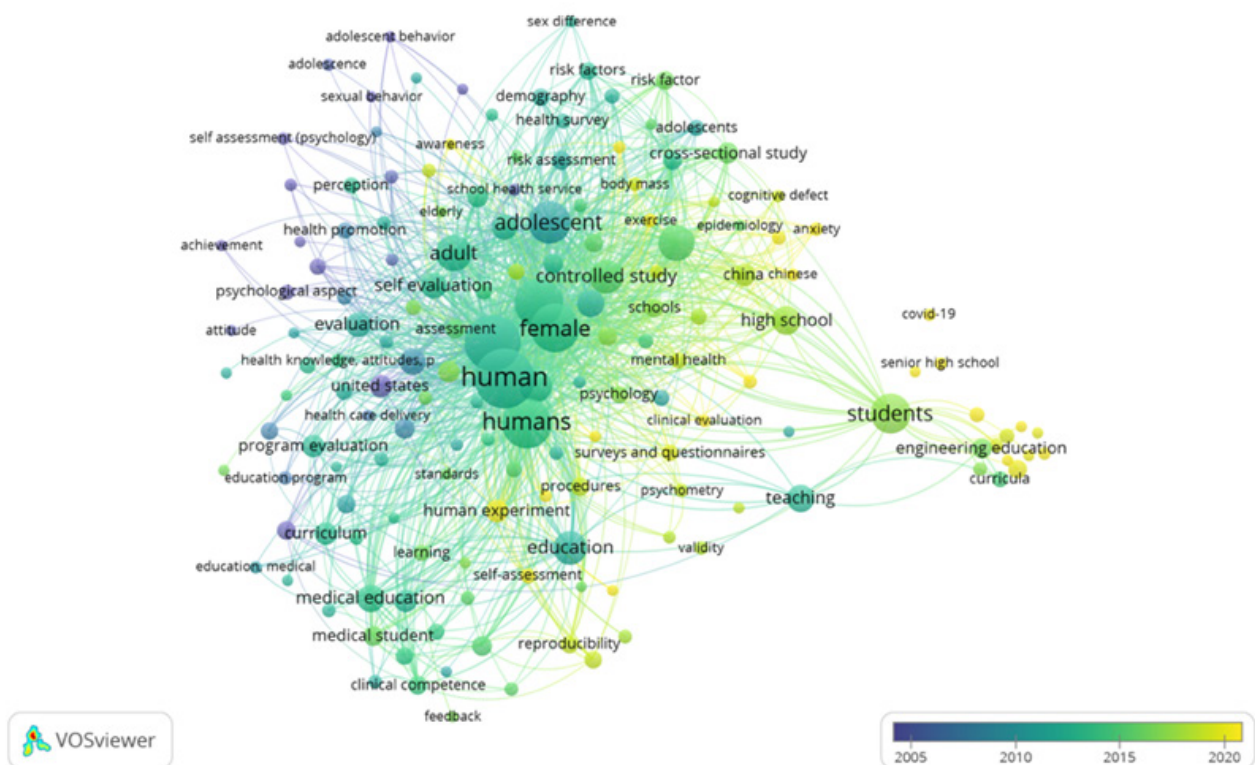


**Figure 3. Bibliometric Shiny visualization**

## Overlay Visualization

Based on the overlay visualization (figure 4), it is evident that research themes on assessment and evaluation in senior high schools have grown significantly over time. Dominant keywords such as students, high school, curricula, and teaching indicate a strong focus on evaluating learning processes, curriculum effectiveness, and the psychological factors of learners, such as anxiety and mental health, within the educational context. The yellow color associated with terms like senior high school, students, and engineering education highlights that these topics are relatively recent and have become central to contemporary educational research, especially in the post-COVID-19 era.

The connections among keywords such as assessment, self-evaluation, surveys, questionnaires, and program evaluation illustrate a shift from traditional assessment models toward approaches grounded in psychometric instruments and self-assessment. This trend is also reflected in the prominence of controlled studies and cross-sectional studies, suggesting that evaluative research is now more empirically rigorous, integrating demographic variables, psychological risk factors, and students' mental health. Hence, the literature network indicates that evaluation in senior high schools extends beyond academic achievement, incorporating holistic aspects of student development. The strong linkages among curriculum, teaching, and education programs with students underscore the importance of adaptive curriculum design in supporting learning. On the right cluster of the visualization, the prominence of engineering education and curricula demonstrates a growing research trajectory toward vocational and STEM education at the senior high school level. This trend aligns with global demands for science and technology literacy and 21st-century career readiness.



**Figure 4.** Overlay visualization from VOSviewer

## DISCUSSION

The cluster 1 (red) illustrates the centrality of programmatic evaluation in advancing medical education by systematic linking learning processes with healthcare delivery outcomes. Studies emphasize the role of program evaluation not only in assessing the effectiveness of curricula but also in refining pedagogical approaches that strengthen clinical competence and professional identity formation among medical students. Feedback mechanisms are particularly highlighted as critical drivers of continuous improvement, enabling learners to bridge gaps between theoretical knowledge and practical application. This body of work underscores the interplay between curriculum design and healthcare system needs, suggesting that educational programs must remain responsive to evolving clinical contexts and patient care standards. The scholarship in this cluster demonstrates that rigorous evaluation practices are indispensable for aligning medical education with broader public health objectives and ensuring sustainable improvements in workforce preparedness.

The cluster 2 (blue) highlights a significant scholarly focus on adolescent health within educational contexts,

where schools serve not only as learning environments but also as critical settings for health assessment and intervention. The prevalence of cross-sectional and controlled study designs reflects the field's commitment to producing robust, evidence-based insights into the multifaceted determinants of youth well-being. Research in this domain consistently emphasizes the interplay between mental health, physical activity, and academic performance, underscoring the importance of school-based health promotion strategies in mitigating risks such as anxiety, cognitive impairments, and maladaptive behaviors. By integrating psychological evaluations with educational outcomes, these studies illuminate how adolescent health directly shapes developmental trajectories, thereby providing essential implications for policy design, curriculum development, and holistic support systems in secondary education.

The emphasis of Cluster 3 (green) on psychological and self-assessment constructs underscores the pivotal role of metacognitive engagement in adolescent learning trajectories, aligning with broader evidence in educational psychology that highlights the interplay between self-awareness, evaluative judgment, and behavioral regulation. Empirical studies have consistently demonstrated that self-assessment practices not only foster critical reflection but also enhance learners' capacity for autonomous goal-setting and adaptive decision-making, both of which are essential for academic resilience during adolescence.<sup>(35)</sup> The senior high school context, cultivating self-evaluation and reflective awareness can be understood as a developmental mechanism that strengthens identity formation while equipping learners with the tools to navigate complex academic and social environments. The application of the law, because judges do not always have to obey written laws, should be the same thing that applies to the world of education.<sup>(48,49)</sup> This cluster's orientation thus situates self-assessment not merely as a pedagogical technique but as a transformative construct that bridges cognitive growth with affective dispositions, offering a comprehensive framework for designing educational interventions that promote lifelong learning and psychological adaptability.

The cluster 4 (Yellow), which emphasizes psychometrics, instrumentation, and methodological validity, underscores the importance of rigorous measurement principles in social science research. Within this domain, scholarly contributions emphasize the dual necessity of theoretical grounding and empirical testing to ensure the robustness of evaluative tools such as surveys, rubrics, and diagnostic instruments. The literature emphasizes that validity and reliability are not static properties of a test, but rather dynamic qualities contingent upon context, population, and purpose, thereby necessitating iterative processes of calibration and reassessment.<sup>(3)</sup> For senior high school educators and researchers, this methodological lens provides both a safeguard and a scaffold, mitigating the risks of measurement bias while enhancing the interpretive accuracy of educational outcomes. The cluster's focus on reproducibility reflects broader scientific imperatives to foster transparency and replicability, aligning educational research with evolving standards of open science and evidence-based practice.

Cluster 5 (purple) underscores the centrality of senior high school education as a focal point for examining contemporary curricular and evaluative practices. The prominence of terms such as "curriculum," "student-centered evaluation," and "teaching" reflects a pedagogical shift toward learner-oriented approaches, where assessment is more progressive, aligned with competencies rather than rote achievement.<sup>(50)</sup> The integration of "engineering education" within this cluster highlights the policy-driven expansion of STEM in secondary curricula, aimed at equipping students with skills relevant to innovation-driven economies.<sup>(51)</sup> The inclusion of "COVID-19" illustrates how the pandemic catalyzed rapid transformations in instructional delivery and assessment, demanding resilience and adaptability from both educators and learners.<sup>(52)</sup> The cluster delineates a nexus where curricular reform, pedagogical strategies, and evaluative frameworks converge, forming the backbone of empirical and theoretical inquiry into high school assessment practices.

The findings of this systematic literature review indicate a clear trajectory in the evolution of assessment and evaluation practices in senior high school education over the past decade. Across the 43 articles included, there is strong evidence of a paradigmatic shift from summative and rote-based models toward competency-based assessment (CBA). This transformation reflects the global educational reform agenda aligned with SDGs 4, which emphasizes inclusive, equitable, and high-quality education for all.<sup>(30)</sup> Studies in diverse contexts, particular in Southeast Asia, demonstrate how assessment practices are increasingly oriented toward higher-order thinking skills, critical reasoning, and reflective judgment.<sup>(3)</sup> Such approaches move beyond mere knowledge recall, fostering transferable skills and learner autonomy. The implementation of CBA remains uneven. Indonesia national policies frequently advocate for assessment models that cultivate creativity and critical thinking, classroom practices often remain rooted in low-order questioning and standardized examinations, a tension also reported in Indonesian contexts where resistance to the Asesmen Kompetensi Minimum (AKM) has been documented.<sup>(1,9)</sup> The documented resistance to the Asesmen Kompetensi Minimum (AKM) further reinforces this interpretation. Teachers do not view the AKM as a meaningful tool to support higher-order learning. Instead, they often see it as an additional burden introduced without enough preparation or systemic support. As a result, teachers return to established practices that emphasize examination performance over developing critical and creative competencies.



Teachers emerge as pivotal agents in shaping the culture of assessment. Multiple studies emphasize that teacher readiness, professional competence, and reflective capacity significantly influence whether innovative assessment strategies are successfully adopted.<sup>(14,15)</sup> Professional development programs have shown promise in equipping teachers with the skills to design authentic and formative assessments, including participatory action research projects and reflective practices embedded in subject-specific instruction. Nonetheless, challenges persist, particularly in decentralized and low-resource systems where opportunities for continuous professional learning are limited. The digital divide further intensifies these challenges, as a significant number of educators are deficient in the necessary infrastructure, technological resources, and pedagogical training essential for the effective implementation of online and hybrid assessment methodologies that were introduced in response to the COVID-19 pandemic.<sup>(6,26)</sup> The deficiency of strategic earmarked financial resources intended for long-term capacity advancement, there is a strong possibility that policy reforms will perpetually be merely aspirational as opposed to effecting transformative change.

In academic discourse, the importance of psychometric accuracy is increasingly emphasized.<sup>(32,41)</sup> The endeavors undertaken have reinforced the methodological underpinnings of assessment research and have facilitated equity across a variety of educational environments. In scenarios where evaluations are associated with significant outcomes such as criteria for graduation or eligibility for advanced academic programs the application of psychometric frameworks becomes especially imperative. Scholars caution that an overreliance on standardized psychometric assessments may compromise the acknowledgment of contextual and cultural dimensions of learning, particularly in multilingual and multicultural contexts. While methodological rigor is of paramount significance, subsequent research initiatives must also incorporate educational values and practices to avert reductionist interpretations of student learning outcomes.<sup>(11,34)</sup>

A notable trend underscores the improvement of evaluative frameworks designed to integrate the multifaceted dimensions of student development. The realm of academic dialogue is progressively claiming that assessment frameworks should incorporate not merely cognitive milestones but also an appraisal of psychosocial and conduct-related elements.<sup>(7,45)</sup> These methodologies promote learner independence by prompting students to assess and control their own development, consequently augmenting motivation and involvement. Academic investigations into the psychological health, general well-being, and risk-associated behaviors of adolescents highlight the complex interaction of psychosocial elements with academic performance.<sup>(31,44)</sup> This aggregation of academic inquiries consistently illustrates the fundamental importance of evaluative frameworks within senior secondary education frameworks to rise above established academic limits, thereby supporting the broader development of students in an era of rapidly transforming social and technological domains.

The COVID-19 pandemic has impacted the integration of technological innovations into assessment frameworks, creating both opportunities and challenges. Studies reveal that students in disadvantaged settings often lacked access to reliable devices, stable internet connectivity, and digital literacy, which in turn affected the fairness and validity of remote assessments.<sup>(33,36)</sup> Teachers, too, faced steep learning curves in adapting their assessment practices to digital environments. In many cases, the absence of adequate infrastructure and professional training intensified stress and undermined instructional quality.<sup>(27)</sup> These findings suggest that digital readiness must be prioritized as a prerequisite for sustainable assessment reforms in the post-pandemic era.

Beyond the confines of pedagogical methodologies, a multitude of research endeavors has broadened the parameters of assessment to encompass institutional efficacy and organizational performance. The application of models such as Data Envelopment Analysis (DEA) situates evaluation within a systemic framework, enabling schools to assess how resources are allocated and utilized.<sup>(13)</sup> Such methodologies expand the evaluative perspective, establishing connections between micro-level instructional strategies and macro-level accountability frameworks. However, although Data Envelopment Analysis (DEA) and analogous quantitative models provide significant insights into efficiency, they ought not to be equated with the notion of educational quality. Efficiency metrics must be interpreted alongside qualitative indicators such as inclusivity, student well-being, and teacher satisfaction to ensure a balanced understanding of institutional performance.

The bibliometric network analysis elucidates that these themes do not exist in a vacuum but are intrinsically linked. For instance, self-assessment methodologies are frequently implemented utilizing validated psychometric tools.<sup>(53,54)</sup> While inquiries into adolescent mental health converge with overarching issues related to curriculum development and instructional methodologies.<sup>(32)</sup> This interconnectedness emphasizes the necessity for integrative frameworks that synthesize pedagogical innovation, psychological support, psychometric validity, and institutional accountability. Whereas comprehensive strategies hold the potential for more sustainable effects.<sup>(26,28)</sup>

National policy frameworks frequently delineate aspirational objectives for competency-based and comprehensive assessment.<sup>(9)</sup> An essential aspect of national education development, which applies universally in various cultures and countries. As national education standards and needs increase, educational development becomes a necessity. This study seeks to investigate the enhancement of teacher quality dimensions to

elevate national education standards, employing a case study approach with mixed methods. The sampling involved random selection from five senior high schools in Lampung, comprising 190 students and 10 teachers. Data collection was carried out through questionnaires, interviews, and document analysis. The analysis of questionnaire data utilized structural equation modeling (SEM). These ambitions are frequently compromised by suboptimal teacher training, a lack of requisite resources, and pervasive cultural. <sup>(20,29)</sup> Bridging this disparity necessitates a bifurcated approach, enhancing teacher capabilities through enduring, contextually relevant professional development and formulating adaptable policies that permit local customization. Policies characterized by excessive prescriptiveness may inadvertently limit innovation. <sup>(55)</sup> The flexible frameworks afford educational institutions the discretion to align assessment methodologies with the specific requirements of their communities. <sup>(17)</sup> These findings highlight several directions for future research. Focus on scalability to determine whether small-scale innovations can be effectively implemented across diverse school systems. <sup>(56)</sup> Sustainability also warrants greater attention, as many reforms falter once external funding or policy support diminishes. <sup>(57)</sup> Cross-cultural validation is essential to ensure that instruments and assessment models are not only reliable but also culturally responsive. <sup>(58,59)</sup> Interdisciplinary collaborations between educators, psychologists, policymakers, and data scientists could further enhance the design of robust, adaptive, and equitable assessment systems capable of addressing the multifaceted challenges of contemporary education. <sup>(60)</sup>

## CONCLUSION

This systematic literature review demonstrated that assessment and evaluation practices in senior high schools are undergoing a significant shift from summative, rote-based models toward competency-based and holistic approaches that emphasize higher-order thinking skills, self-assessment, and reflective learning. The systematic literature review of articles from scopus revealed that teacher agency and professional development are central to the successful implementation of these reforms, while systemic challenges such as policy-practice misalignment, insufficient training, and the digital divide remain persistent barriers. The integration of psychometric rigor and validated instruments has strengthened the methodological foundation of assessment research, yet the need for cultural responsiveness and contextual adaptation persists. The expansion of assessment to include psychosocial well-being and institutional efficiency underscores its multidimensional character, linking classroom pedagogy with broader educational quality and equity. By synthesizing pedagogical, psychological, and organizational perspectives, this study contributes to a more comprehensive understanding of assessment as both a measure and a driver of educational improvement. Future research should focus on scalability, sustainability, and cross-context validation to ensure that assessment models remain adaptable, equitable, and responsive to the evolving demands of 21st-century education.

## REFERENCES

1. Estrera EB. A Perception-Based Curricular Review on the K to 12 HUMSS Strand Curriculum. *ije*. 2020;8(4):25-44.
2. Priyadi A, Kuswanto H. The Android-based comic of Gajah Mungkur Dam: Improving mathematical representation and critical thinking abilities. *J Technol Sci Educ*. 2023;13(1):116.
3. Kriswantoro K, Kartowagiran B, Rohaeti E. A Critical Thinking Assessment Model Integrated with Science Process Skills on Chemistry for Senior High School. *EUROPEAN J ED RES*. 2021;10(1):285-98.
4. Asaju K. Achieving the Sustainable Development Goals (SDGs) and the Intricacies and Dynamics of Development Administration. *CSI*. 2022;2(2):173.
5. Limbong RJ, Fitri OR. The Human Rights Framework for Covid-19 Pandemic Recovery. *CSI*. 2023;3(2):158.
6. Rong R, Chen W, Dai Z, Gu J, Chen W, Zhou Y, et al. Improvement of the management of mental well-being and empathy in Chinese medical students: a randomized controlled study. *BMC Med Educ*. 2021;21(1):378. Available from: <https://bmcmmededuc.biomedcentral.com/articles/10.1186/s12909-021-02813-6>
7. Wulandari P, Asrowi A, Djono D. Preliminary Study on the Development of Infographic-Based Islamic Education Teaching Materials to Improve Soft Skills for High School Students in Surakarta. *PAIJPENDISLAM*. 2022;19(1):29-40.
8. Tawil M, Said MA, Suryansari K. Authentic assessment development science to assess student competency. 2023;11(2):194-206.

9. Dacholfany MI, Suseno N, Syofyan H, Fadli MR. Educational management in improving the quality of teachers in senior high schools. *IJERE*. 2024;13(5):3100.
10. Waty ERK, Nengsih YK, Handrianto C, Rahman MA. The quality of teacher-made summative tests for Islamic education subject teachers in Palembang, Indonesia. *CP*. 2024;43(1):186-97.
11. Alforque JM, Jovero MB, Bodanio C, Ecle N, Picardal JP. Content Analysis of Bioenergetics' Conceptual Presentation in Senior High School Biology Textbooks. *J Bio Bio Edu*. 2024;16(2):363-72.
12. Juwaini J, Rijal S, Basir @ Ahmad A, Desky H, Tamtowi M, Mawardi M, et al. Ibn Miskawaih's Ethical Philosophy and Its Relevance to Moral Education in Indonesian Secondary Schools. *jip*. 2025;13(1):695-720.
13. Durmuşçelebi M. Examination of Students' Academic Motivation, Research Concerns and Research Competency Levels during the Education Period. *ujer*. 2018;6(10):2115-24.
14. Keshmiri F, Jambarsang S, Mehrparvar AH. Effective components of teachers' professionalism in viewpoints of various stakeholders. *Journal of Education and Health Promotion*. 2023;12(1). Available from: [https://journals.lww.com/10.4103/jehp.jehp\\_1565\\_21](https://journals.lww.com/10.4103/jehp.jehp_1565_21)
15. Pineda H, Lorca A, Cortes S, Gador S, Mangompit RM, Pacaldo FJ, et al. Development And Evaluation of A Professional Development Program on Designing Participatory Action Research Projects for Basic Education Teachers. *AE*. 2022;161-84. Available from: <http://ae.fl.kpi.ua/article/view/266663>
16. Bosisio J. A research landscape on customer co-creation value: a systematic literature network analysis. *Ital J Mark*. 2024;2024(3):339-68.
17. Jopling R, Attah D, Abas M, Adde KS, Goldsmith K, Bere T, et al. African Youth in Mind - Protocol of a Pilot feasibility trial of a brief psychological Intervention for older adolescents with depression delivered through senior high schools in Navrongo, Ghana. Al-Khateeb Z, editor. *PLoS ONE*. 2025;20(4):e0319462.
18. Mohzana M, Muh. Fahrurrozi MuhF, Abdurrosyidin R. AR. Improving the Performance of Administrative Staff: An Analysis of Evaluation and Its Implications in Mataram City High Schools. *EDUPIJ*. 2025;15(1). Available from: <https://www.edupij.com/index/arsiv/76/482/improving-the-performance-of-administrative-staff-an-analysis-of-evaluation-and-its-implications-in-mataram-city-high-schools>
19. Sarfika R, Malini H, Wicaksana AL, Wenny BP, Saifudin IMMY. Cross cultural adaptation and psychometric evaluation of the Indonesian version of the depression anxiety stress scales for youth (IDASS-Y). *Heliyon*. 2024;10(19):e38830.
20. Haque TF, Knudsen JE, You J, Hui A, Djaladat H, Ma R, et al. Competency in Robotic Surgery: Standard Setting for Robotic Suturing Using Objective Assessment and Expert Evaluation. *Journal of Surgical Education*. 2024;81(3):422-30.
21. Taningrum N, Kriswanto ES, Pambudi AF, Yulianto WD. Improving Critical Thinking Skills Using Animated Videos Based on Problem-Based Learning. *Retos*. 2024;57:692-6.
22. Insani M, Haenilah EY, Hariri H, Sinaga RM. Audio-visual-based history learning media materials about human life in the literary age. *EduLearn*. 2023;17(3):398-407.
23. Agustin MG, Cabansag MGS. Disaster readiness and risk reduction management module using Kolb's model. *IJERE*. 2023;12(2):703.
24. Sumarno S, Gimin G, Noviana E. The Comparison Analysis of Teacher's Learning Effectiveness in Realizing Entrepreneurship Skills of Students. *INT J INSTRUCTION*. 2023;16(2):679-702.
25. De Oliveira AM, Santos JLF, Capellini SA. Reading comprehension performance of elementary and senior high school students. *Front Educ*. 2023;8:1086040.
26. Liswati TW, Mustaji ., Hariyati N, Uulaa RFR. Actualization of Principal's Visionary Leadership in Improving

Teacher Skills in Managing Learning Innovations in Senior High Schools. *JESR.* 2023;13(1):45.

27. Sari IJ, Pongsophon P, Vongsangnak W, Pimthong P, Pitiporntapin S. The development of molecular genetics concept test for senior high school students using Rasch analysis. *IJERE.* 2022;11(4):1687.

28. Soeharto S, Csapó B. Assessing Indonesian student inductive reasoning: Rasch analysis. *Thinking Skills and Creativity.* 2022;46:101132.

29. Herman H, Shara AM, Silalahi TF, Sherly S, Julyanthry J. Teachers' Attitude towards Minimum Competency Assessment at Sultan Agung Senior High School in Pematangsiantar, Indonesia. *JCT.* 2022;11(2):1.

30. Fatchurahman M, Setiawan MA, Karyanti K. Evaluation of the CSE-UCLA model on the performance of school counselor in Indonesia. *P Sci Edu.* 2022;56(2):561-72.

31. Yang Z, Chen W, Jin M, Chen W, Chen L, Zhou X. Analysis of factors influencing casual sexual behavior among male college students in Zhejiang Province, China. Wang Z, editor. *PLoS ONE.* 2021;16(5):e0250703. Available from: <https://dx.plos.org/10.1371/journal.pone.0250703>

32. Zhu X, Shek DTL. Subjective Outcome Evaluation of a Positive Youth Development Program in Mainland China. *Research on Social Work Practice.* 2021;31(3):285-97.

33. Chen Y, Jiang M, Kesten O. An empirical evaluation of Chinese college admissions reforms through a natural experiment. *Proc Natl Acad Sci USA.* 2020;117(50):31696-705.

34. Mamolo LA, Sugano SGC. Self-perceived and actual competencies of senior high school students in General Mathematics. Hui SKF, editor. *Cogent Education.* 2020;7(1):1779505.

35. ZulfiKar Z, Hidayah N, Triyono T, Hitipeuw I. Development Study of Academic Resilience Scale for Gifted Young Scientists Education. *Journal for the Education of Gifted Young Scientists.* 2020;8(1):343-59.

36. Laksmi OD, Chung MH, Liao YM, Chang PC. Multidimensional Scale of Perceived Social Support in Indonesian adolescent disaster survivors: A psychometric evaluation. Oppong Asante K, editor. *PLoS ONE.* 2020;15(3):e0229958.

37. Yansen D, Putri RII, Zulkardi Z, Fatimah S. Developing Pisa-Like Mathematics Problems On Uncertainty and Data Using Asian Games Football Context. *J Math Educ.* 2019;10(1):37-46.

38. Tjabolo SA, Otaya LG. The Evaluation of Islamic Education Teachers Performance. *Jurn Pend Islam.* 2019;5(1):25-38.

39. Hadarah H, Tulhikmah R. Epistemology Philosophy Development of Assessment Instruments Student Chemistry Learning Outcomes. *Journal for the Education of Gifted Young Scientists.* 2019;7(3):627-41.

40. Ad'hiya E, Laksono EW. Development and Validation of an Integrated Assessment Instrument to Assess Students' Analytical Thinking Skills in Chemical Literacy. *INT J INSTRUCTION.* 2018;11(4):241-56.

41. Spence JP, Buddenbaum JL, Bice PJ, Welch JL, Carroll AE. Independent investigator incubator (I3): a comprehensive mentorship program to jumpstart productive research careers for junior faculty. *BMC Med Educ.* 2018;18(1):186.

42. Nizar H, Putri RII, Zulkardi Z. Developing Pisa-Like Mathematics Problem Using The 2018 Asian Games Football and Table Tennis Context. *J Math Educ.* 2018;9(2):183-94.

43. Tully V, Murphy D, Fioratou E, Chaudhuri A, Shaw J, Davey P. Learning from errors: assessing final year medical students' reflection on safety improvement, five year cohort study. *BMC Med Educ.* 2018;18(1):57.

44. Shen L, Yang L, Zhang J, Zhang M. Benefits of expressive writing in reducing test anxiety: A randomized controlled trial in Chinese samples. Moitra E, editor. *PLoS ONE.* 2018;13(2):e0191779. Available from: <https://dx.plos.org/10.1371/journal.pone.0191779>



45. Kruzhkova OV, Vorobyeva IV, Bruner TI, Krivoshchekova MS. The value sphere of native and newcomer youth in their subjective assessment of the environment of a megalopolis. *Psych Rus.* 2015;8(4):139-54.
46. Wibowo S, Dimiyati K, Absori A, Wardiono K, Ramon TM, Budiono A, et al. Islamic nomocracy: from the perspectives of Indonesia, Spain and Russia. *LJIH.* 2023;31(1):91-111.
47. Wardiono K. Prophetic: An Epistemological Offer for Legal Studies. *JTL.* 2019;1(1):17-41.
48. Iriani D, Mahfiana L, Ningrum E. Political Power and Politic Ethical Integrity Of Election Organizers (A Study Of The Development Of The Philosophy Of Science). *JTL.* 2022;4(1):56-70.
49. Budiono A, Yuspin W, Nurani SS, Fairuzzaman F, Pradnyawan SWA, Sari SD. The Anglo-Saxon System of Common Law and the Development of the Legal System in Indonesia. *WSEAS Transactions on Systems.* 2023;22:207-13.
50. Alotaibi H, S. Al-Khalifa H, AlSaeed D. Teaching Programming to Students with Vision Impairment: Impact of Tactile Teaching Strategies on Student's Achievements and Perceptions. *Sustainability.* 2020;12(13):5320.
51. Campbell C, Hobbs L, Xu L, McKinnon J, Speldewinde C. Girls in STEM: Addressing SDG 4 in Context. *Sustainability.* 2022;14(9):4897.
52. Martin S, Dillon J, Higgins P, Peters C, Scott W. Divergent Evolution in Education for Sustainable Development Policy in the United Kingdom: Current Status, Best Practice, and Opportunities for the Future. *Sustainability.* 2013;5(4):1522-44.
53. Kocoglu-Tanyer D, Yorulmaz Demir DS. Development and psychometric evaluation of the vaccine knowledge test for childhood vaccinations: using Rasch analysis. *BMC Public Health.* 2025;25(1):2449.
54. Jarl M, Escher C, Harbut P, Conte H, Nilsson U. Psychometric evaluation of a structured assessment tool for nurse anesthetists' non-technical skills. *BMC Med Educ.* 2025;25(1):718.
55. Ma C, Rattanaburi K, Gao C. Burned out parents, educated kids? Navigating "Juan" culture: Decision-making-process and education in Thailand. *International Journal of Educational Research Open.* 2025;9:100458.
56. Tawil M, Said MA, Suryansari K. Authentic assessment development science to assess student competency. 2023;11(2):194-206.
57. Scholz F, Börner N, Schust SA, Schardey J, Kühn F, Renz B, et al. Focus on patient perspectives in climate action policies for healthcare. A German survey analysis on what patients are willing to do. *Front Public Health.* 2024;12:1477313.
58. Chen M, Dutt AS, Nair R. Cross-cultural adaptation and validation of the Chinese version of Skills and Needs Inventories in Functional Behavior Assessments and Interventions (SNI-FBAI-CN). *Behavioral Interventions.* 2024;39(4):e2038.
59. Nicastrì M, Dincer D'Alessandro H, Anderson K, Ciferri M, Cavalcanti L, Greco A, et al. Cross-Cultural Adaptation and Validation of the Listening Inventory for Education-Revised in Italian. *Audiology Research.* 2024;14(5):822-39.
60. Zhang H, Wu L, Wang J. Research on innovative models of industry-education integration to promote sustainable education in agricultural machinery majors. *Sustainable Futures.* 2025;10:101072. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S2666188825006367>

#### FINANCING

No financing.

#### CONFLICT OF INTEREST

The authors declare that there is no conflict of interest.

#### **AUTHORSHIP CONTRIBUTION**

*Conceptualization:* Mukhamad Zulianto, Arief Budiono, Irmadatus Sholekhah, Arum Putri Rahayu, Turdialiev Mukhammad Ali Polatjon ogli.

*Software:* Irmadatus Sholekhah, Turdialiev Mukhammad Ali Polatjon ogli.

*Writing - original draft:* Mukhamad Zulianto, Arief Budiono, Irmadatus Sholekhah.

*Writing - review and editing:* Mukhamad Zulianto, Arief Budiono, Turdialiev Mukhammad Ali Polatjon ogli.