MUDALLA : PROCEEDING INTERNATIONAL CONFERENCE ON ARABIC LANGUAGE
ISSN : 2807-8780
(MUDALLA) المؤتمر الدولي للغة العربية وآدابها وتعليمها
قسم الأدب العربي جامعة مالانج الحكومية

Are International-Class Students Future Ready? A Benchmarking Study of 21st-Century Competencies in Arabic Language Education

Muh. Sabilar Rosyad
Universitas Islam Negeri Sunan Kalijaga Yogyakarta, Indonesia
muhammad.rosyad@uin-suka.ac.id

Abstract

This study benchmarks the 21st-century competencies of students enrolled in an International Class Program in Arabic Language Education to evaluate their readiness for future educational and professional demands. Using a cross-sectional benchmarking survey, data were collected from 25 undergraduate students through a validated questionnaire based on the Partnership for 21st Century Learning (P21) Framework, encompassing ten competency domains. Descriptive statistics, competency ranking, gap analysis, and a competency benchmark matrix were employed to identify strengths and curriculum priorities. The findings indicate that students demonstrated a high overall level of 21st-century competencies (M = 4.05, SD = 0.49), suggesting that the international-class environment effectively supports future-ready graduate development. Communication, Collaboration, and Global Competence emerged as the strongest competencies, whereas Leadership, Digital Literacy, and Creativity represented the largest competency gaps. The study extends existing research by transforming competency assessment into an evidence-based benchmarking framework that informs curriculum evaluation and continuous quality improvement. The proposed Competency Benchmark Matrix offers practical guidance for curriculum enhancement and provides a baseline for longitudinal benchmarking across cohorts and institutions, thereby strengthening future-ready Arabic language teacher education.

Keywords Competency Assessment; Curriculum Evaluation; Digital Literacy; P21 Framework; Teacher Education

مستخلص البحث

تهدف هذه الدراسة إلى قياس كفايات القرن الحادي والعشرين لدى طلبة البرنامج الدولي في تعليم اللغة العربية، بهدف التعرف إلى مدى جاهزيتهم لمتطلبات الدراسة والعمل في المستقبل. واعتمدت الدراسة منهج المسح المقطعي المقارن، وشملت 25 طالبًا في مرحلة البكالوريوس، جُمعت بياناتهم من خلال استبانة محكَّمة بُنيت على إطار الشراكة لمهارات القرن الحادي والعشرين(P21 Framework) ، الذي يغطي عشرة مجالات للكفايات. ولتحليل البيانات، استُخدمت الإحصاءات الوصفية، وترتيب الكفايات، وتحليل الفجوات، إلى جانب مصفوفة المقارنة المعيارية للكفايات؛ لتحديد جوانب التميز وأولويات تطوير المنهج .أظهرت النتائج أن الطلبة يتمتعون بمستوى مرتفع من كفايات القرن الحادي والعشرين بصفة عامة (M = 4.05, SD = 0.49) مما يعكس فاعلية بيئة البرنامج الدولي في إعداد خريجين يمتلكون المهارات اللازمة للمستقبل. وجاءت مهارات التواصل، والعمل التعاوني، والكفاية العالمية في مقدمة الكفايات الأكثر قوة، بينما ظهرت القيادة، والثقافة الرقمية، والإبداع بوصفها

المجالات الأكثر حاجة إلى التطوير .وتقدم هذه الدراسة إسهامًا جديدًا من خلال تحويل تقويم الكفايات إلى إطار مرجعي قائم على الأدلة يمكن الاستفادة منه في تقويم المناهج وتحسينها بصورة مستمرة. كما توفر مصفوفة المقارنة المعيارية للكفايات أداة عملية لتطوير المناهج، وتمثل نقطة انطلاق لإجراء مقارنات معيارية مستقبلية بين الدفعات والمؤسسات التعليمية، بما يدعم إعداد معلمي اللغة العربية القادرين على مواكبة متطلبات المستقبل.

كلمات أساسية: تقويم الكفايات؛ تقويم المناهج؛ الثقافة الرقمية؛ إطار مهارات القرن الحادي والعشرين (P21)؛ إعداد معلمي اللغة العربية

Introduction

The accelerating pace of technological innovation, globalisation, and digital transformation has fundamentally redefined the competencies expected of university graduates in the twenty-first century. Advances in artificial intelligence, automation, big data, and digital communication have transformed labour markets, requiring graduates to possess not only disciplinary expertise but also higher-order thinking, adaptability, collaboration, and lifelong learning capabilities. This transformation is further reinforced by the shift from Industry 4.0 toward Society 5.0, which promotes the integration of advanced technologies with human-centred values to address increasingly complex societal challenges (Alves et al., 2023; Ghobakhloo et al., 2026; Martini et al., 2024). Consequently, higher education institutions are under growing pressure to prepare future-ready graduates who can respond effectively to dynamic professional environments while contributing responsibly to sustainable social development. International organisations, including UNESCO, the Organisation for Economic Co-operation and Development (OECD), the World Economic Forum (WEF), the Partnership for 21st Century Learning (P21), and the Assessment and Teaching of 21st Century Skills (ATC21S), consistently emphasise that universities should cultivate critical thinking, creativity, communication, collaboration, digital literacy, and lifelong learning as essential competencies for success in an uncertain future (Ehlers, 2020; Lee et al., 2025; Thornhill-Miller et al., 2023). These competencies have therefore become widely recognised as key indicators of graduate quality and institutional effectiveness in contemporary higher education.

Within this global educational transformation, teacher education occupies a particularly strategic position because future teachers are expected not only to master subject knowledge but also to facilitate learning that equips students with competencies required for the twenty-first century. The P21 framework conceptualises these competencies through three interconnected domains, namely learning and innovation skills, information, media and technology skills, and life and career skills (Eisler, 2002; Guo & Woulfin, 2016), all of which are increasingly

regarded as fundamental outcomes of teacher education programmes. Previous studies have consistently demonstrated that teachers who possess strong critical thinking, creativity, communication, collaboration, and digital competencies are more capable of implementing innovative pedagogical practices, integrating educational technologies, facilitating learner-centred instruction, and adapting to rapidly changing educational contexts (Fitria, 2025; Huang et al., 2024; Lestari & Kurnia, 2023; Nurhayati et al., 2025). Nevertheless, cultivating these competencies remains a significant challenge, particularly in language teacher education, where prospective teachers must simultaneously develop linguistic expertise, intercultural awareness, technological proficiency, and professional adaptability. While twenty-first-century competencies have become central to teacher education worldwide, Arabic language teacher education has only recently begun to integrate these competencies systematically into curriculum and instructional practices, indicating the need for more comprehensive evaluations of students' preparedness beyond conventional measures of language proficiency and academic achievement.

The growing internationalisation of higher education has encouraged universities to establish international-class programmes as strategic initiatives for producing globally competitive graduates. These programmes generally provide bilingual or multilingual instruction, internationally benchmarked curricula, greater exposure to multicultural learning environments, and opportunities for cross-cultural academic collaboration. Such learning environments are expected to strengthen communication skills, intercultural competence, global awareness, collaborative abilities, and adaptability, thereby supporting the development of future-ready graduates. Within Arabic language teacher education, international-class programmes offer distinctive opportunities for students to engage with authentic language use, international academic resources, multicultural interactions, and digitally mediated learning experiences that potentially enhance their professional readiness (Bahtiar, 2025; Rosyad, 2025). However, these expectations remain largely theoretical unless supported by empirical evidence. Participation in an international-class programme does not necessarily guarantee balanced development across all dimensions of twenty-first-century competencies, as students may demonstrate excellence in communication and intercultural interaction while remaining less prepared in areas such as critical thinking, leadership, creativity, or digital literacy (Kaltsum et al., 2025; Pamessangi, 2019; Utsman & Rosyad, 2023). Therefore, systematically evaluating the competency profile of international-class students is essential for determining whether these programmes effectively fulfil their intended educational objectives.

As higher education increasingly emphasises evidence-based quality assurance, benchmarking has emerged as an important approach to curriculum evaluation and continuous programme improvement (Asif, 2015; Hämäläinen et al., 2002; Tam, 2014). Unlike conventional descriptive assessments that merely report average competency scores, benchmarking establishes systematic reference points that enable institutions to identify strengths, diagnose developmental needs,

monitor educational progress, and evaluate curriculum effectiveness over time. Competency benchmarking is particularly relevant for teacher education because it provides a comprehensive profile of students' preparedness across multiple competency domains while generating evidence that can inform strategic curriculum development. Rather than merely reporting students' competency levels, benchmarking provides baseline evidence for curriculum evaluation and future improvement (Henderson-Smart et al., 2006; Tasopoulou & Tsiotras, 2017). Through a competency benchmark profile, programme administrators and curriculum developers can identify priority areas requiring pedagogical intervention, evaluate the effectiveness of educational innovations, and establish measurable indicators for future curriculum enhancement. Consequently, benchmarking offers a more meaningful framework for evaluating international-class programmes than descriptive reporting alone because it directly connects assessment findings with institutional decision-making and continuous quality improvement.

Despite increasing scholarly interest in internationalisation, teacher education, and Arabic language learning, several important research gaps remain. Existing studies in Arabic language education have predominantly examined language proficiency, academic achievement, instructional innovation, learning motivation, technology integration, and students' perceptions, while relatively little attention has been devoted to evaluating comprehensive twenty-first-century competencies (Adawiyah & Qibthiyah, 2026; Farisha et al., 2023; Hizbullah et al., 2025; Kosim et al., 2025; Safrudin et al., 2024). Furthermore, although international-class programmes have become increasingly common in higher education, most previous investigations have focused on students' academic performance or international learning experiences rather than assessing whether these programmes genuinely foster future-ready competencies. Within Arabic language teacher education, research examining future readiness through internationally recognised competency frameworks remains particularly limited, leaving insufficient empirical evidence regarding the preparedness of prospective Arabic language teachers for contemporary educational challenges (AbdelWahhab, 2026; Widyasari, 2025; Yan Mei et al., 2025). More importantly, to the best of the authors' knowledge, no previous study has developed a competency benchmark profile for international-class students in Arabic Language Education that can serve as baseline evidence for curriculum evaluation and continuous programme improvement. Addressing these gaps is therefore expected to contribute both theoretically by extending discussions on future-ready teacher education and practically by informing evidence-based curriculum development.

In response to these limitations, this study adopts a competency benchmarking perspective to evaluate the twenty-first-century competencies of international-class students enrolled in an Arabic Language Education programme. Rather than treating competency assessment as a purely descriptive exercise, the

study seeks to generate a comprehensive competency profile that can function as a benchmark for evaluating curriculum effectiveness and guiding future programme development. Specifically, the study aims to map students' overall level of twenty-first-century competencies based on the P21 framework, identify competency domains that represent students' greatest strengths and those requiring further improvement, and develop a competency benchmark profile that serves as baseline evidence for curriculum evaluation. The resulting findings are expected to support curriculum developers, programme administrators, and teacher educators in designing targeted educational interventions that strengthen future-ready competencies while enhancing the quality and international competitiveness of Arabic language teacher education.

Accordingly, this study addresses two fundamental questions that guide the competency benchmarking process. First, to what extent do international-class students demonstrate twenty-first-century competencies? Second, which competencies are strongest and weakest among international-class students? By answering these questions, the study seeks to move beyond conventional evaluations of academic performance and language proficiency toward a more comprehensive understanding of graduate preparedness. The competency benchmark profile generated through this research is expected to provide an empirical baseline for future curriculum evaluation, facilitate comparisons across cohorts or institutions, and support continuous improvement initiatives aimed at preparing future-ready Arabic language teachers capable of responding effectively to the educational demands of the twenty-first century.

Method

This study employed a cross-sectional benchmarking survey to examine the 21st-century competencies of students enrolled in the International Class Program of the Arabic Language Education Department (Chen & Wu, 2006; Wang & Cheng, 2020; Zangirolami-Raimundo et al., 2018). A cross-sectional design was selected because it enables researchers to capture students' competency profiles at a single point in time, providing a systematic snapshot of their readiness for future educational and professional demands. Unlike conventional descriptive surveys that primarily report average scores (Coughlan et al., 2009; Shkedi, 2004), the benchmarking approach was adopted to generate a competency benchmark profile that serves as baseline evidence for curriculum evaluation and continuous programme improvement. By comparing competency levels across multiple domains, benchmarking facilitates the identification of students' strengths, developmental priorities, and competency gaps that can inform evidence-based curriculum enhancement. Accordingly, this study positions competency assessment as a strategic tool for programme evaluation rather than merely a descriptive measurement of student performance.

The study was conducted in the International Class Program of the Undergraduate Arabic Language Education Department, UIN Sunan Kalijaga Yogyakarta, Indonesia. The programme represents one of the department's internationalisation initiatives, integrating bilingual instruction, internationally oriented learning resources, intercultural learning experiences, and technology-enhanced instruction to prepare globally competent Arabic language teachers. As such, the programme provides an appropriate context for assessing students' readiness to meet the competencies expected of future educators in the twenty-first century.

The participants consisted of 25 undergraduate students enrolled in the International Class Program from Semester 2 to Semester 4 during the 2025/2026 academic year. Because the number of students in the programme was relatively limited, a total sampling technique was employed, whereby all eligible students were invited to participate (Levy & Lemeshow, 2013; Petersen et al., 2005). Restricting participation to students who had completed at least one semester ensured that respondents had experienced the academic environment and learning practices implemented within the international-class curriculum. Participation was voluntary, and all respondents completed the questionnaire in its entirety.

The research instrument was developed based on the Partnership for 21st Century Learning (P21) Framework, which provides a comprehensive conceptualisation of competencies required for future-ready graduates. The questionnaire comprised 30 items distributed across ten competency domains: Critical Thinking, Creativity, Communication, Collaboration, Information Literacy, Digital Literacy, Self-directed Learning, Leadership, Adaptability, and Global Competence. Each domain was represented by three statements measured using a five-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree) (Robinson, 2023; Tanujaya et al., 2022). The questionnaire items were adapted from established instruments on twenty-first-century competencies and teacher education, followed by contextual modification to reflect the characteristics of Arabic language teacher education.

To ensure instrument quality, the questionnaire underwent a two-stage validation process. First, two experts in Arabic language education and curriculum studies evaluated the instrument to establish content validity, focusing on item relevance, clarity, and contextual appropriateness. Their feedback resulted in minor revisions to improve wording and content representation. Second, the revised questionnaire was piloted with students outside the research sample to examine item clarity and internal consistency. Reliability analysis using Cronbach's alpha indicated satisfactory reliability for the overall instrument and all competency domains, with coefficients exceeding the recommended threshold of 0.70, confirming the instrument's suitability for data collection (Cronbach, 1951; Santos, 1999).

Data were collected between June and July 2026 through an online questionnaire distributed via the programme's official communication channels. Before completing the survey, participants received detailed information regarding the purpose of the study, voluntary participation, confidentiality, anonymity, and the exclusive use of the collected data for academic research. Electronic informed consent was obtained from every participant before accessing the questionnaire. No personally identifiable information was collected, and participants were informed that they could withdraw from the study at any stage without academic consequences.

Data analysis was conducted in five sequential stages to generate a comprehensive competency benchmark profile (Avram, 2025). First, an Overall Competency Score was calculated by averaging all questionnaire responses to determine students' overall level of twenty-first-century competencies. Second, Domain Mean Analysis was performed to examine competency levels across the ten domains. Third, the domains were organised through Competency Ranking to identify the strongest and weakest competencies demonstrated by the participants. Fourth, a Competency Gap Analysis was conducted by comparing each domain's empirical mean with the maximum attainable score, thereby identifying competency areas requiring greater curricular attention. Finally, the findings were synthesised into a Competency Benchmark Matrix, categorising each competency according to its benchmark level and corresponding curriculum priority. This five-stage analytical procedure constitutes the principal methodological contribution of the study by transforming descriptive survey findings into evidence that can support curriculum evaluation and continuous improvement.

Results and Discussion

Benchmark Profile across Competency Domains

The first research question examined the overall level of twenty-first-century competencies demonstrated by international-class students in the Arabic Language Education programme. Descriptive analysis revealed an overall mean score of 4.05 (SD = 0.49), indicating that students generally demonstrated a high level of twenty-first-century competencies according to the predetermined interpretation criteria. This finding suggests that the international class learning environment has contributed positively to students' preparedness for future educational and professional contexts.

Although the overall competency profile reflects a favourable level of future readiness, the distribution of scores across competency domains indicates that competency development was not uniform. Several competencies achieved excellent performance, while others remained at a moderate level, suggesting that some areas require further curricular attention. Therefore, the overall findings indicate that international-class students possess a solid foundation of future-ready

competencies; however, targeted curriculum enhancement remains necessary to achieve balanced competency development across all domains.

Table 1. Overall 21st-Century Competency Profile

Indicator	Value
Number of participants	25
Overall Mean	4.05
Standard Deviation	0.49
Competency Level	High

To provide a more comprehensive understanding of student's competency profiles, descriptive statistics were calculated for each competency domain. Table 4 presents the mean score, standard deviation, ranking, benchmark level, and competency gap for all ten domains.

The findings indicate considerable variation across competency domains. Communication obtained the highest mean score (M = 4.48), followed by Collaboration (M = 4.39) and Global Competence (M = 4.27). These competencies consistently achieved high benchmark levels, suggesting that students benefited from the collaborative, bilingual, and internationally oriented learning environment.

Conversely, Leadership recorded the lowest mean score (M = 3.62), followed by Digital Literacy (M = 3.73) and Creativity (M = 3.82). Although these scores remain above the midpoint of the measurement scale, they indicate relatively weaker competency development compared with other domains.

Table 2. Benchmark Profile across Competency Domains

Competency Domain	Mean	SD	Rank	Benchmark Level	Gap (5.00 – Mean)
Communication	4.48	0.42	1	Excellent	0.52
Collaboration	4.39	0.46	2	Excellent	0.61
Global Competence	4.27	0.48	3	High	0.73
Critical Thinking	4.16	0.51	4	High	0.84
Adaptability	4.09	0.49	5	High	0.91
Self-directed Learning	3.97	0.56	6	Moderate	1.03
Information Literacy	3.91	0.58	7	Moderate	1.09
Creativity	3.82	0.61	8	Moderate	1.18
Digital Literacy	3.73	0.64	9	Moderate	1.27
Leadership	3.62	0.67	10	Moderate	1.38

Strongest and Weakest Competencies

The second research question sought to identify the strongest and weakest competency domains among international-class students. The results demonstrate that Communication, Collaboration, and Global Competence emerged as the three strongest competencies. Their consistently high mean scores indicate that students were generally confident in expressing ideas effectively, collaborating with peers,

and interacting within multicultural academic environments. These findings are consistent with the characteristics of international-class programmes, which emphasise bilingual instruction, collaborative learning, academic discussion, and intercultural engagement.

In contrast, Leadership represented the weakest competency domain, followed by Digital Literacy and Creativity. The relatively lower performance in these competencies suggests that although students have developed strong interpersonal and global communication skills, they have fewer opportunities to exercise leadership roles, initiate innovative learning activities, or engage critically with emerging digital technologies beyond routine academic use. Furthermore, the relatively large competency gaps observed in these domains indicate priority areas for future curriculum development. Therefore, the findings suggest that international-class programmes have been successful in strengthening competencies associated with global engagement but remain less effective in cultivating competencies related to educational leadership, creative problem solving, and advanced digital capability.

To transform descriptive findings into evidence for curriculum evaluation, the competency scores were synthesised into a Competency Benchmark Matrix. Rather than presenting competency levels in isolation, the matrix classifies each domain according to its benchmark category and corresponding curriculum priority. This matrix constitutes the principal outcome of the benchmarking approach because it provides programme administrators with an evidence-based reference for curriculum review and continuous quality improvement.

The matrix reveals three strategic competency clusters. The first cluster comprises Communication and Collaboration, which reached excellent benchmark levels and therefore should be maintained through continued implementation of collaborative and bilingual learning practices. The second cluster includes Global Competence, Critical Thinking, and Adaptability, all of which demonstrated high performance but would benefit from further strengthening through inquiry-based learning, authentic problem-solving tasks, and increased international academic engagement. Finally, the third cluster consists of Self-directed Learning, Information Literacy, Creativity, Digital Literacy, and Leadership, which require greater curricular attention because they represent the largest competency gaps. These findings provide a clear roadmap for curriculum enhancement by identifying priority competencies that require systematic intervention.

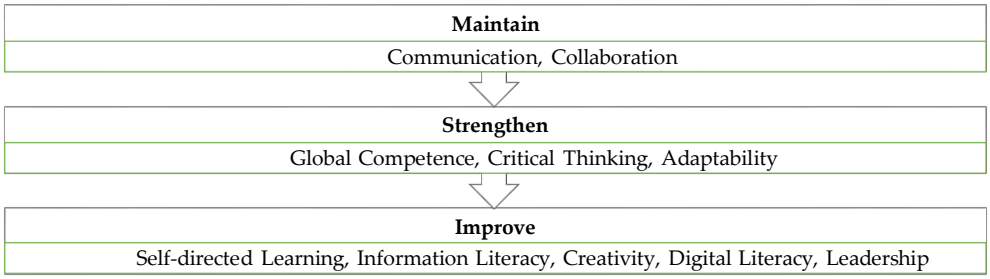

Figure 1. Competency Benchmark Matrix

Based on the benchmarking analysis, the competency profile demonstrates that international-class students are generally globally prepared to participate in multilingual, collaborative, and intercultural learning environments. Nevertheless, the unequal distribution of competencies indicates that future curriculum development should prioritise strengthening leadership, digital literacy, creativity, and self-directed learning to ensure that prospective Arabic language teachers are comprehensively prepared to meet the educational demands of the twenty-first century. Accordingly, the Competency Benchmark Matrix provides baseline evidence that can support systematic curriculum evaluation, monitor future programme improvements, and facilitate longitudinal benchmarking across student cohorts.

Are International-Class Students Future Ready?

The findings of this study provide a comprehensive benchmark of the twenty-first-century competencies demonstrated by international-class students in the Arabic Language Education programme. The discussion therefore focuses on four interconnected issues: students' overall readiness for the twenty-first century, the competencies that constitute the programme's strengths, the competencies requiring further development, and the contribution of competency benchmarking to evidence-based curriculum evaluation.

The overall competency profile indicates that international-class students demonstrated a high level of twenty-first-century competencies (M = 4.05, SD = 0.49), suggesting that they are generally well prepared to respond to the academic and professional demands of contemporary education. This finding supports the growing consensus that higher education should prepare graduates who possess not only disciplinary expertise but also transferable competencies such as communication, collaboration, critical thinking, adaptability, and global awareness. From the perspective of the P21 Framework (Eisler, 2002; Guo & Woulfin, 2016), the competency profile observed in this study reflects substantial progress toward developing future-ready graduates capable of functioning effectively in increasingly complex and interconnected educational environments.

The findings are broadly consistent with previous studies reporting that international learning environments contribute positively to the development of communication skills, intercultural competence, collaborative learning, and learner adaptability (Luong et al., 2024; Purbasari et al., 2025; Thinh, 2024; Vaishnav, 2025). Similarly, research in teacher education has consistently shown that authentic learning experiences, multilingual interaction, and technology-supported learning environments facilitate the acquisition of competencies required for twenty-first-century teaching (McCarthy et al., 2023; Rosyad et al., 2023). Nevertheless, the

present findings also indicate that future readiness should not be interpreted as a uniform construct. Although the overall competency level was high, considerable variation remained across competency domains, suggesting that students' preparedness is characterised by differentiated strengths rather than balanced competency development. Therefore, future readiness should be understood as a multidimensional construct requiring continuous curriculum support rather than a final educational outcome achieved simply through participation in an international-class programme.

Importantly, these findings reinforce the argument that evaluating future readiness should extend beyond traditional indicators such as grade point average or language proficiency. Graduates may achieve excellent academic performance while still demonstrating uneven development in competencies essential for professional success. Consequently, competency benchmarking provides a more comprehensive understanding of graduate preparedness by examining multiple dimensions that collectively shape future professional capability.

Patterns of 21st-Century Competency Development among International-Class Students

Among the ten competency domains, Communication, Collaboration, and Global Competence emerged as the strongest competencies. These findings are unsurprising given the distinctive characteristics of the international-class programme, which emphasises bilingual instruction, collaborative learning, international academic engagement, and multicultural interaction. Students are routinely required to communicate ideas in more than one language, participate in collaborative projects, present academic work publicly, and interact with diverse learning resources. Such learning experiences naturally encourage the development of interpersonal communication and collaborative problem-solving skills (Ningsih et al., 2025; Tuimebayeva et al., 2024).

The high level of global competence also reflects the programme's international orientation. Exposure to multilingual academic materials, cross-cultural perspectives, and internationally referenced curricula appears to strengthen students' awareness of global issues while fostering openness toward cultural diversity. This finding aligns with theories of internationalisation in higher education, which argue that international learning environments should cultivate graduates capable of working across linguistic, cultural, and geographical boundaries (Gregersen-Hermans & Lauridsen, 2021; Larsen & Larsen, 2016). For prospective Arabic language teachers, global competence represents an increasingly important professional attribute because language education is inherently connected with intercultural understanding and global communication.

Critical Thinking and Adaptability also achieved relatively high benchmark levels, indicating that students possess a satisfactory capacity to analyse information, respond to changing learning situations, and adjust to unfamiliar academic environments. Although these competencies did not reach the excellent

category, their relatively strong performance suggests that the programme has successfully integrated learning activities encouraging analytical reasoning and flexible learning behaviours. Collectively, these findings demonstrate that the international-class curriculum effectively nurtures competencies associated with communication, collaboration, and intercultural engagement, thereby establishing a strong foundation for preparing globally competent Arabic language educators.

Despite the generally positive findings, several competency domains demonstrated comparatively lower performance, particularly Leadership, Digital Literacy, Creativity, Information Literacy, and Self-directed Learning. Although these competencies remained within the moderate category rather than indicating poor performance, they exhibited the largest competency gaps relative to the ideal benchmark. This finding suggests that international-class participation alone is insufficient to ensure balanced development across all dimensions of future readiness.

Leadership emerged as the weakest competency, indicating that students may have limited opportunities to assume leadership roles within academic activities. Many university learning environments continue to position students primarily as participants rather than organisers, decision-makers, or project leaders (Duignan, 2012; Menon, 2005). Consequently, students may develop strong collaborative abilities without simultaneously cultivating leadership competencies such as strategic decision-making, initiative, conflict management, and team coordination. Given that graduates of Arabic Language Education programmes are expected to become educational leaders, curriculum designers should provide more structured opportunities for leadership development through project management, peer mentoring, student-led seminars, community engagement, and collaborative teaching practices (Taha, 2017).

Digital Literacy also requires further attention despite students' familiarity with digital technologies. The distinction between routine technology use and advanced digital literacy is particularly important. Effective digital literacy involves evaluating information critically, selecting appropriate digital resources, integrating educational technologies pedagogically, maintaining digital ethics, and using emerging technologies such as artificial intelligence responsibly (Adamakis & Rachiotis, 2025; Falloon, 2020). Similarly, relatively lower scores in Creativity and Self-directed Learning suggest that students may still depend heavily on instructor-directed learning and conventional classroom activities. These findings highlight the importance of expanding inquiry-based learning, design thinking, project-based learning, reflective practice, and authentic problem-solving activities that encourage innovation and learner autonomy. Therefore, strengthening these competencies is essential if future Arabic language teachers are expected to lead educational innovation rather than merely implement existing instructional practices.

Benchmarking Implications for Curriculum Development

The principal contribution of this study lies in demonstrating how competency benchmarking can support evidence-based curriculum evaluation within Arabic language teacher education. Unlike conventional descriptive studies that simply report competency levels, the competency benchmark profile developed in this research provides practical guidance for curriculum planning by translating empirical findings into strategic curriculum priorities. This approach enables programme administrators to identify competencies that should be maintained, strengthened, or improved through targeted educational interventions.

The benchmark profile suggests that curriculum enhancement should adopt a differentiated rather than uniform strategy. Competencies already reaching excellent levels should continue to be reinforced through collaborative learning, bilingual instruction, and international academic engagement. Competencies within the high category require systematic strengthening through inquiry-oriented learning, authentic assessment, and increased opportunities for critical reflection. Meanwhile, competencies categorised as moderate should become priority targets for curriculum innovation because they represent the largest developmental gaps affecting graduates' future readiness.

Table 3. Curriculum Benchmark Recommendations Based on Competency Profile

Benchmark Finding	**Recommended Curriculum Action**
Communication (Excellent)	Maintain bilingual academic discussions, presentations, and international academic communication activities.
Collaboration (Excellent)	Expand collaborative projects, peer learning, and cross-cultural teamwork with partner institutions.
Global Competence (High)	Increase international exposure through virtual exchange programmes, intercultural dialogue, and collaborative international research.
Critical Thinking (High)	Strengthen inquiry-based learning, problem-based learning, and evidence-based academic writing.
Adaptability (High)	Integrate experiential learning and authentic teaching practice in diverse educational settings.
Self-directed Learning (Moderate)	Encourage reflective learning portfolios, independent research projects, and personalised learning plans.
Information Literacy (Moderate)	Improve students' ability to evaluate, synthesise, and ethically use academic information from international sources.
Creativity (Moderate)	Incorporate design thinking, innovation projects, and creative instructional material development.
Digital Literacy (Moderate)	Integrate AI literacy, digital pedagogy, educational technology design, and digital ethics across the curriculum.
Leadership (Moderate)	Develop structured leadership experiences through student-led projects, mentoring, academic organisations, and collaborative microteaching.

Overall, the competency benchmark profile developed in this study offers a practical framework for continuous curriculum improvement in Arabic language

teacher education. Rather than functioning solely as an assessment instrument, benchmarking becomes a strategic quality assurance mechanism that enables departments to monitor competency development longitudinally, evaluate the effectiveness of curriculum reforms, and compare student cohorts over time (Moriarty, 2011; Ossiannilsson, 2023). Consequently, this study extends the application of benchmarking beyond institutional performance measurement by demonstrating its value as an evidence-based approach for designing future-ready curricula that respond to the evolving demands of twenty-first-century language teacher education.

Conclusion

This study benchmarked the twenty-first-century competencies of international-class students in an Arabic Language Education programme using the P21 Framework. The findings indicate that the students demonstrated an overall high level of future-ready competencies, suggesting that the international-class programme has effectively fostered communication, collaboration, global competence, and critical thinking. Nevertheless, the competency profile revealed uneven development across domains, with leadership, digital literacy, creativity, information literacy, and self-directed learning requiring greater curricular attention. The Competency Benchmark Matrix provides an evidence-based framework for identifying curriculum priorities, enabling programme administrators to distinguish competencies that should be maintained, strengthened, or improved. Accordingly, competency benchmarking offers a practical approach to continuous curriculum evaluation and quality enhancement in Arabic language teacher education.

This study is limited by its small sample size and focus on a single institution, which may restrict the generalisability of the findings. Future research should involve multiple universities, larger and more diverse samples, and longitudinal or mixed-methods designs to examine the development of twenty-first-century competencies over time and across different international higher education contexts.

Acknowledgement

The author gratefully acknowledges the Arabic Language Education Department, Faculty of Tarbiyah and Education, UIN Sunan Kalijaga Yogyakarta, for its institutional support and for providing a conducive academic environment that facilitated the successful completion of this research. The author also sincerely thanks the students of the International Class Program in Arabic Language Education who voluntarily participated in this study and generously shared their time, experiences, and perspectives. Their valuable contributions were essential to

the completion of this research and the development of evidence-based insights into 21st-century competencies in Arabic language education.

References

AbdelWahhab, F. R. (2026). *Exploring the Perspectives of AFL Teachers on Teacher Development Needs, with Reference to the Level of Professional Experience*.

Adamakis, M., & Rachiotis, T. (2025). Artificial Intelligence in Higher Education: A State-of-the-Art Overview of Pedagogical Integrity, Artificial Intelligence Literacy, and Policy Integration. In *Encyclopedia* (Vol. 5, Issue 4, p. 180). https://doi.org/10.3390/encyclopedia5040180

Adawiyah, R., & Qibthiyah, L. (2026). Research Trends in Machine Learning for Arabic Language Studies: A Bibliometric Analysis. *Al-Afkar, Journal For Islamic Studies*, *9*(3), 373–381.

Alves, J., Lima, T. M., & Gaspar, P. D. (2023). Is Industry 5.0 a Human-Centred Approach? A Systematic Review. In *Processes* (Vol. 11, Issue 1, p. 193). https://doi.org/10.3390/pr11010193

Asif, M. (2015). Determining improvement needs in higher education benchmarking. *Benchmarking: An International Journal*, *22*(1), 56–74. https://doi.org/10.1108/BIJ-02-2013-0025

Avram, K. A. (2025). Introducing the Staged Narrative Analysis: A Comprehensive Framework to Analyze Multi-Layered and Complex Narrative Data. *Forum: Qualitative Social Research/Qualitative Sozialforschung*, *26*(1).

Bahtiar, M. F. (2025). *Imersi Bahasa Arab pada Program Kelas Internasional Mahasiswa Prodi PBA S1 Fakultas Ilmu Tarbiyah dan Keguruan UIN Sunan Kalijaga Yogyakarta*. UIN SUNAN KALIJAGA YOGYAKARTA.

Chen, Z.-G., & Wu, K. H. (2006). Comparison of Benchmarking Methods with and without a Survey Error Model. *International Statistical Review*, *74*(3), 285–304. https://doi.org/https://doi.org/10.1111/j.1751-5823.2006.tb00297.x

Coughlan, M., Cronin, P., & Ryan, F. (2009). Survey research: Process and limitations. *International Journal of Therapy and Rehabilitation*, *16*(1), 9–15. https://doi.org/10.12968/ijtr.2009.16.1.37935

Cronbach, L. J. (1951). Coefficient Alpha and the Internal Structure of Tests. *Psychometrika*, *16*(3), 297–334. https://doi.org/DOI: 10.1007/BF02310555

Duignan, P. (2012). *Educational leadership: Together creating ethical learning environments*. Cambridge University Press.

Ehlers, U.-D. (2020). *Future skills: The future of learning and higher education*. BoD–Books on Demand.

Eisler, R. (2002). Partnership Education For the 21st Century. *Encounter*, *15*(3).

Falloon, G. (2020). From digital literacy to digital competence: the teacher digital competency (TDC) framework. *Educational Technology Research and Development*, *68*(5), 2449–2472. https://doi.org/10.1007/s11423-020-09767-4

Farisha, F., Fahri, M., & Al-Ghifari, R. A. (2023). Analysis Of Research Trends In

Arabic Language Education Using A Bibliometric Approach In Indonesian National Journals/Analisis Tren Penelitian Pendidikan Bahasa Arab Menggunakan Pendekatan Bibliometrik Pada Jurnal Nasional Indonesia. *Lughawiyah: Journal of Arabic Education and Linguistics*, *5*(2), 177–190.

Fitria, I. (2025). Guru Sebagai Fasilitator Digital: Membangun Kompetensi Pedagogik. *Manajemen Pendidikan Mi/Sd: Berbasis Teknologi Dan Neurosains Dalam Kurikulum Merdeka*, *233*.

Ghobakhloo, M., Mahdiraji, H. A., Iranmanesh, M., & Jafari-Sadeghi, V. (2026). From Industry 4.0 digital manufacturing to Industry 5.0 digital society: A roadmap toward human-centric, sustainable, and resilient production. *Information Systems Frontiers.*, 1–33. https://doi.org/10.1007/s10796-024-10476-z

Gregersen-Hermans, J., & Lauridsen, K. M. (2021). *Internationalising programmes in higher education: An educational development perspective*. Routledge.

Guo, J., & Woulfin, S. (2016). Twenty-First Century Creativity: An Investigation of How the Partnership for 21st Century Instructional Framework Reflects the Principles of Creativity. *Roeper Review*, *38*(3), 153–161. https://doi.org/10.1080/02783193.2016.1183741

Hämäläinen, K., Hämäläinen, K., Jessen, A. D., Kaartinen-Koutaniemi, M., & Kristoffersen, D. (2002). Benchmarking in the improvement of higher education. *ENQA Workshop Reports*, *2*.

Henderson-Smart, C., Winning, T., Gerzina, T., King, S., & Hyde, S. (2006). Benchmarking learning and teaching: developing a method. *Quality Assurance in Education*, *14*(2), 143–155. https://doi.org/10.1108/09684880610662024

Hizbullah, I., Fauzi, A., Antarsyach, A. A., & Fathurrahman, M. A. (2025). Analyse Applied Linguistics Research Trends in Google Scholar Using a Bibliometric Approach with VOSviewer. *Al-Fashahah Journal*, *5*(2). https://doi.org/10.59562/al-fashahah.v5i2.76703

Huang, R., Adarkwah, M. A., Liu, M., Hu, Y., Zhuang, R., & Chang, T. (2024). Digital Pedagogy for Sustainable Education Transformation: Enhancing Learner-Centred Learning in the Digital Era. *Frontiers of Digital Education*, *1*(4), 279–294. https://doi.org/https://doi.org/10.1007/s44366-024-0031-x

Kaltsum, U., Nuruddin, A., Sodik, A. J., Fu'adah, S., & Rosyad, M. S. (2025). Problematika Pembelajaran Bahasa Arab di Indonesia Berbasis Analisis Fishbone. *Muhadasah: Jurnal Pendidikan Bahasa Arab*, *7*(2), 227–250. https://doi.org/10.51339/muhad.v7i2.4529

Kosim, N., Fauzia, E. L., Safrudin, R., Fu'adah, S., & Siregar, S. D. P. (2025). A Bibliometric Study of Research Trends in Arabic Language Bilingualism (1986-2024). *Arabiyatuna: Jurnal Bahasa Arab*, *9*(1), 67–94.

Larsen, M. L., & Larsen, M. A. (2016). *Internationalization of higher education*. Springer. https://doi.org/10.1057/978-1-137-53345-6

Lee, N. A. A., Tazijan, F. N., Adam, A. F. M., Ikhsanudin, I., & Bakar, R. A. (2025). Comparative analysis of digital literacy and 21st-century skills among

university graduates in Malaysia and Indonesia: The role of collaboration, critical thinking, communication, and creativity. *Journal of Nusantara Studies (JONUS)*, *10*(1), 166–191. https://doi.org/10.24200/jonus.vol10iss1pp166-191

Lestari, D. I., & Kurnia, H. (2023). Implementasi model pembelajaran inovatif untuk meningkatkan kompetensi profesional guru di era digital. *JPG: Jurnal Pendidikan Guru*, *4*(3), 205–222. https://doi.org/10.32832/jpg.v4i3.14252

Levy, P. S., & Lemeshow, S. (2013). *Sampling of populations: methods and applications*. John Wiley & Sons.

Luong, P. M., Tran, L. T., Nguyen, H. T., Tran, Y. T. H., Dang, G. H., & Vu, T. Van. (2024). Intercultural adaptability development for students in internationalisation at home programmes in higher education in China and Korea. *Higher Education, Skills and Work-Based Learning*, *15*(2), 421–440. https://doi.org/10.1108/HESWBL-10-2023-0284

Martini, B., Bellisario, D., & Coletti, P. (2024). Human-Centered and Sustainable Artificial Intelligence in Industry 5.0: Challenges and Perspectives. In *Sustainability* (Vol. 16, Issue 13, p. 5448). https://doi.org/10.3390/su16135448

McCarthy, A. M., Maor, D., McConney, A., & Cavanaugh, C. (2023). Digital transformation in education: Critical components for leaders of system change. *Social Sciences and Humanities Open*, *8*(1), 100479. https://doi.org/10.1016/j.ssaho.2023.100479

Menon, M. E. (2005). Students' Views Regarding their Participation in University Governance: Implications for Distributed Leadership in Higher Education. *Tertiary Education and Management*, *11*(2), 167–182. https://doi.org/10.1007/s11233-005-0686-x

Moriarty, J. P. (2011). A theory of benchmarking. *Benchmarking: An International Journal*, *18*(4), 588–611. https://doi.org/10.1108/14635771111147650

Ningsih, T. Z., Aman, A., Nasrulloh, A., Ofianto, O., Erniwati, E., Asri, Z., Judijanto, L., & Firza, F. (2025). Enhancing communication and collaboration skills through discovery, cooperative and problem-based learning models in Social Studies education. *Cogent Education*, *12*(1), 2500110. https://doi.org/10.1080/2331186X.2025.2500110

Nurhayati, S., Septikasari, D., Judijanto, L., Susanto, D., Tarrapa, S., Sudadi, S., Setiyana, R., Willdahlia, A. G., Ramli, A., & Zamroni, Z. (2025). *Paradigma baru dalam pendidikan abad 21*. PT Green Pustaka Indonesia.

Ossiannilsson, E. S. I. (2023). *Benchmarking: A Method for Quality Assessment and Enhancement in Higher Education BT - Learning, Design, and Technology: An International Compendium of Theory, Research, Practice, and Policy* (J. M. Spector, B. B. Lockee, & M. D. Childress (eds.); pp. 2149–2167). Springer International Publishing. https://doi.org/10.1007/978-3-319-17461-7_52

Pamessangi, A. A. (2019). Analisis kesulitan belajar bahasa Arab mahasiswa program studi pendidikan bahasa Arab IAIN Palopo. *AL IBRAH: Journal of Arabic Language Education*, *2*(1), 11–24. https://doi.org/10.24256/jale.v2i1.1206

Petersen, L., Minkkinen, P., & Esbensen, K. H. (2005). Representative sampling for

reliable data analysis: Theory of Sampling. *Chemometrics and Intelligent Laboratory Systems*, *77*(1), 261–277. https://doi.org/https://doi.org/10.1016/j.chemolab.2004.09.013

Purbasari, I., Yusuf, M., Subagya, null, Marmoah, S., & Alali, A. A. (2025). Development of a community-based social collaborative e-learning model for adaptability of higher education students in Indonesia. *Frontiers in Education*, *Volume 10-2025*.

Robinson, J. (2023). *Likert Scale BT - Encyclopedia of Quality of Life and Well-Being Research* (F. Maggino (ed.); pp. 3917–3918). Springer International Publishing. https://doi.org/10.1007/978-3-031-17299-1_1654

Rosyad, M. S. (2025). Student-Centered Arabic Teaching Materials in Internasional Undergraduate Programs: Empirical Evidence from UIN Sunan Kalijaga Yogyakarta. *PROSIDING KONIPBSA: Konferensi Nasional Inovasi Pembelajaran Bahasa Dan Sastra Arab*, *5*(1), 217–232.

Rosyad, M. S., Adam, F. M., & Bahruddin, U. (2023). *The Natural Situation and Its Effective Role in Designing Arabic Listening Skills Teaching Materials for Non-Native Speakers*. *6*(3), 586–604. https://doi.org/https://doi.org/10.18860/ijazarabi.v6i3.19531

Safrudin, R., Nandang, A., Siregar, S. D. P., Musthafa, I., Fauzi, M. F., Alby, M. H. F., & Suyono, S. (2024). Development of Arabic Language Learning Research: A Bibliometric Study on Scopus (2009-2024). *Al-Ta'rib: Jurnal Ilmiah Program Studi Pendidikan Bahasa Arab IAIN Palangka Raya*, *12*(2), 321–338.

Santos, J. R. A. (1999). Cronbach's alpha: A tool for assessing the reliability of scales. *The Journal of Extension*, *37*(2), 15. https://doi.org/10.66752/1077-5315.7392

Shkedi, A. (2004). Narrative survey: A methodology for studying multiple populations. *Narrative Inquiry*, *14*(1), 87–111. https://doi.org/10.1075/ni.14.1.05shk

Taha, H. (2017). Arabic language teacher education. *Applied Linguistics in the Middle East and North Africa*, 269–287.

Tam, M. (2014). Outcomes-based approach to quality assessment and curriculum improvement in higher education. *Quality Assurance in Education*, *22*(2), 158–168. https://doi.org/10.1108/QAE-09-2011-0059

Tanujaya, B., Prahmana, R. C. I., & Mumu, J. (2022). Likert scale in social sciences research: Problems and difficulties. *FWU Journal of Social Sciences*, *16*(4), 89–101. https://doi.org/10.51709/19951272/Winter2022/7

Tasopoulou, K., & Tsiotras, G. (2017). Benchmarking towards excellence in higher education. *Benchmarking: An International Journal*, *24*(3), 617–634. https://doi.org/10.1108/BIJ-03-2016-0036

Thinh, M. P. (2024). Adapting teaching practices and enhancing intercultural competence: student teachers' experiences in an international exchange program. *Journal for Multicultural Education*, *19*(1), 43–57. https://doi.org/10.1108/JME-06-2024-0077

Thornhill-Miller, B., Camarda, A., Mercier, M., Burkhardt, J.-M., Morisseau, T., Bourgeois-Bougrine, S., Vinchon, F., El Hayek, S., Augereau-Landais, M., Mourey, F., Feybesse, C., Sundquist, D., & Lubart, T. (2023). Creativity, Critical Thinking, Communication, and Collaboration: Assessment, Certification, and Promotion of 21st Century Skills for the Future of Work and Education. In *Journal of Intelligence* (Vol. 11, Issue 3, p. 54). https://doi.org/10.3390/jintelligence11030054

Tuimebayeva, G., Shagrayeva, B., Kerimbayeva, K., Shertayeva, N., Bitemirova, A., & Abdurazova, P. (2024). *Developing Multilingual Competence in Future Educators: Approaches, Challenges, and Best Practices*. *6*(1). https://doi.org/doi:10.1515/edu-2024-0020

Utsman, M., & Rosyad, M. S. (2023). Fenomena Rendahnya Minat Menguasai Keterampilan Membaca serta Solusinya dari Sudut Pandang Siswa dan Guru. *LUGHATI: Jurnal Pendidikan Bahasa Arab*, *1*(01), 1–16.

Vaishnav, P. B. (2025). The role of intercultural competence in effective cross-border collaboration among higher education institutions. In *International Academic Transformations and Cross-Border Collaborations* (pp. 195–218). IGI Global Scientific Publishing. https://doi.org/10.4018/979-8-3373-0508-0.ch009

Wang, X., & Cheng, Z. (2020). Cross-Sectional Studies: Strengths, Weaknesses, and Recommendations. *Chest*, *158*(1, Supplement), S65–S71. https://doi.org/https://doi.org/10.1016/j.chest.2020.03.012

Widyasari, M. P. (2025). *Dilema Guru: Antara Kompetensi dan Tuntunan Kurikulum*. Indonesia Emas Group.

Yan Mei, S., Ju, S. Y., Romli, T. R. M., & Adam, Z. B. (2025). Enhancing employability: a study on professional skills development among Arabic language graduates in Malaysia. *Education + Training*, *67*(7–8), 802–820. https://doi.org/10.1108/ET-10-2024-0485

Zangirolami-Raimundo, J., Echeimberg, J. de O., & Leone, C. (2018). Research methodology topics: Cross-sectional studies. *Journal of Human Growth and Development*, *28*(3), 356–360. https://doi.org/10.7322/jhgd.152198