

Assessing and Developing Universal Competencies of Pedagogical College Students

Fatima Mohammed Almalki¹, Abdulghani Ali Al-Hattami²

Bahrain Teachers College University of Bahrain

DOI: <https://doi.org/10.47772/IJRISS.2026.100700477>

Received: 23 July 2026; Accepted: 28 July 2026; Published: 05 August 2026

ABSTRACT

This study primarily aimed to develop and validate the Universal Competency Assessment Tool (UCAT) as an assessment instrument aligned with the Education and Training Quality Authority's (BQA) 21st-century skills framework. Using data from 623 undergraduate teacher-education students at Bahrain Teachers College, UCAT's measurement quality was evaluated through reliability and structural validity evidence, drawing on both self-perception indicators and performance-based scoring. Results supported the intended eight-domain structure: exploratory factor analysis using principal-axis extraction with oblimin rotation indicated strong sampling adequacy ($KMO = 0.923$) and a significant Bartlett's test ($\chi^2(780) = 5828.163$, $p < .001$), and confirmatory factor analysis further demonstrated excellent fit for an eight-factor correlated model ($\chi^2(712) = 711.157$, $CFI = 1.000$, $TLI = 1.000$, $RMSEA = 0.000$, $SRMR = 0.030$). Performance scoring showed strong interrater reliability across domains ($ICC(2,1) = 0.714\text{--}0.880$; weighted $\kappa = 0.713\text{--}0.879$). As a secondary aim, UCAT was used to estimate students' competency levels and to quantify self-performance discrepancies, revealing consistently higher self-perceived scores than performance-based scores ($d_z = 0.310\text{--}0.762$). Finally, the study explored whether students' reported exposure to selected instructional practices was associated with performance outcomes; these models explained only small proportions of variance ($R^2 = 0.005\text{--}0.023$). Overall, the findings provide evidence that UCAT yields structurally valid and dependable scores for assessing universal competencies in teacher education and underscore the value of triangulating self-report with performance evidence in competency-based assessment.

Keywords: Universal Competencies; 21st-Century Skills; Performance-Based Assessment; Teacher Education; Factor Analysis; Interrater Reliability

INTRODUCTION

The rapid transformation of global education systems in response to the demands of the Fourth Industrial Revolution has shifted the focus of teacher education from mere content mastery toward the cultivation of universal competencies, a multidimensional set of skills necessary for personal, academic, and professional success in the 21st century. Referred to as soft skills, transversal skills, or future skills, these competencies encompass a broad range of abilities including critical thinking, creativity, collaboration, communication, digital literacy, global awareness, emotional intelligence, adaptability, and resilience (OECD, 2018; UNESCO, 2015). In teacher education, these skills are not ancillary but foundational; they empower educators to navigate dynamic classroom environments, foster inclusive and innovative pedagogical practices, and support holistic student development.

Despite widespread international endorsement of competency-based education (CBE) frameworks (Mulder, 2014; González & Wagenaar, 2008), significant gaps remain in operationalizing, assessing, and contextualizing universal competencies within pedagogical institutions, particularly in non-Western contexts. Studies have emphasized the lack of standardized, validated tools for measuring these competencies (Shavelson et al., 2018), inconsistencies in curricular alignment (Voogt & Roblin, 2012), and the limited empirical data capturing longitudinal growth among teacher candidates. Moreover, existing assessment models often overlook cultural and institutional specificities, creating challenges in adapting global frameworks to localized education systems

(Hofstede, 1984).

In the Kingdom of Bahrain, the Education and Training Quality Authority (BQA) has taken a decisive step in addressing these challenges by embedding 21st-century skills within its national Schools Reviews Framework (BQA, 2024). This framework identifies eight core competencies, critical thinking and problem solving, collaboration and communication, creativity and innovation, information literacy, digital literacy, global and cultural awareness, adaptability and resilience, and emotional intelligence, as essential for learner success in a knowledge-based and diversified economy envisioned under Bahrain's Economic Vision 2030. While these skills are widely recognized across international benchmarks, their classification by the BQA as universal competencies affirms their relevance for all disciplines and levels of education.

However, there is limited empirical evidence assessing how these universal competencies are being cultivated among future educators in Bahrain. Bahrain Teachers College (BTC) is expected to align its curricula and pedagogical approaches with BQA standards, yet systematic investigations into the actual competency levels of pedagogical students and the effectiveness of instructional strategies remain scarce. Additionally, the extent to which existing assessment practices reflect or support the development of these universal competencies has not been adequately explored. Addressing these knowledge gaps is critical to ensuring that future educators are not only subject-matter proficient but also equipped with the universal skills needed to thrive in modern educational contexts.

This study primarily aims to develop and validate the Universal Competency Assessment Tool (UCAT) as an assessment instrument aligned with the Education and Training Quality Authority's (BQA) 21st-century skills framework. As a secondary aim, the study uses UCAT to estimate students' competency levels and to quantify the discrepancy between self-perceived and performance-based competency scores. Finally, the study explores whether students' reported exposure to selected instructional practices is associated with UCAT performance scores.

LITERATURE REVIEW

Universal competencies, often equated with 21st-century skills or soft skills, encompass a broad spectrum of cognitive, interpersonal, and intrapersonal capabilities necessary for success in dynamic educational environments (OECD, 2018). These include critical thinking, creativity, collaboration, communication, adaptability, digital literacy, global awareness, and emotional intelligence. According to the World Economic Forum (2020), such competencies are not only valuable but essential in preparing future educators to navigate increasingly complex classroom scenarios.

In the context of Bahrain's educational landscape, the Education and Training Quality Authority (BQA) has emphasized the integration of 21st-century skills within its Review Frameworks (BQA, 2024) to align with the Fourth Industrial Revolution and Bahrain's Economic Vision 2030. The BQA framework aligns closely with global models, particularly the OECD's Learning Compass (2019) and UNESCO's transversal competencies framework (2016). These competencies are deemed essential for students to navigate the complexities of the modern world and are categorized as follows:

Critical Thinking and Problem Solving: This skill involves the ability to analyze information objectively, evaluate various perspectives, and devise effective solutions to complex problems. In Bahrain's educational framework, fostering critical thinking is crucial for preparing students to address real-world challenges.

Collaboration and Communication: Emphasizing teamwork, this competency requires students to work effectively with others, share ideas clearly, and engage in constructive dialogue. Bahrain's focus on this skill aims to enhance students' interpersonal abilities, essential for both academic and professional success.

Creativity and Innovation: Encouraging students to think outside the box, this skill involves generating original ideas and embracing innovative approaches. Incorporating creativity into the curriculum is vital for Bahrain to cultivate a generation capable of driving progress and adapting to change.

Information Literacy: This competency entails the ability to locate, evaluate, and effectively use information from various sources. In an era of information overload, Bahrain's educational system prioritizes teaching students to discern credible sources and apply information judiciously.

Digital Literacy: With the increasing integration of technology, digital literacy involves understanding and utilizing digital tools responsibly. Bahrain's emphasis on this skill ensures that students are prepared for the technological demands of the contemporary workplace.

Global and Cultural Awareness: This skill fosters an appreciation for cultural diversity and global perspectives. By promoting this competency, Bahrain aims to prepare students to operate in an interconnected world, respecting and valuing different cultures.

Adaptability and Resilience: Focusing on the ability to adjust to new situations and recover from setbacks, this skill is essential for personal and professional growth. Bahrain's educational initiatives strive to build these traits to help students navigate the uncertainties of the future.

Emotional Intelligence: Involving self-awareness, empathy, and social skills, emotional intelligence is crucial for personal development and effective interpersonal interactions. Bahrain's focus on this competency aims to enhance students' emotional well-being and social harmony.

By embedding these 21st-century skills into its educational review frameworks, Bahrain is proactively preparing its students to meet the evolving demands of the global economy and society. This strategic alignment not only addresses current educational needs but also contributes to the realization of Bahrain's Economic Vision 2030, fostering a knowledge-based and diversified economy.

The assessment and development of universal competencies in teacher education are grounded in robust theoretical frameworks, most notably Competency-Based Education (CBE) and socio-constructivist learning theory. CBE has become a cornerstone in the reform of higher education, emphasizing that the focus of instructional systems should shift from time-bound curricula to outcome-based, skill-centered models. According to Mulder (2014), CBE ensures that educational success is determined by demonstrable mastery of specific competencies, such as critical thinking, digital literacy, and collaboration, rather than the passive accrual of academic credits. This approach holds particular value in teacher education, where graduates must possess both content knowledge and the capacity to apply that knowledge in diverse and unpredictable classroom settings.

Vygotsky's (1978) socio-constructivist theory complements CBE by highlighting the role of guided learning, social interaction, and contextualized problem-solving in fostering higher-order skills such as critical thinking and collaboration. These theoretical paradigms underpin many recent studies advocating experiential learning, peer engagement, and interdisciplinary projects to promote competency development (Voogt & Roblin, 2012; Clarke & Hollingsworth, 2017). Voogt and Roblin (2012) argue that 21st-century skills require instructional approaches that are interdisciplinary, inquiry-based, and rooted in real-world problem solving. Their framework calls for the transformation of teacher education programs into environments where learning is experiential, collaborative, and adaptive mirroring the complex realities teachers will face in professional practice. Similarly, Clarke and Hollingsworth (2017) present an Interconnected Model of Professional Growth, which illustrates how reflection, experimentation, and external feedback interact to support sustained competency acquisition among both pre-service and in-service teachers.

A growing body of empirical research has sought to evaluate the extent to which pre-service teachers demonstrate mastery of universal competencies, with findings revealing a complex and often uneven landscape. While many studies underscore the increasing recognition of these competencies within teacher preparation programs, there remains a persistent gap between policy aspirations and actual student outcomes. For example, Nguyen et al. (2020) conducted a large-scale, cross-national study across teacher education institutions in East and Southeast Asia to assess the self-reported and performance-based proficiency of pedagogical students in various 21st-century competencies. The results indicated relatively high levels of digital literacy, particularly in the use of educational technologies, digital content creation, and engagement with learning management systems. However,

the same cohort exhibited notable deficiencies in creativity, resilience, and global cultural awareness. These gaps were attributed in part to rigid curricula, limited exposure to real-world problem-solving, and a predominant focus on technical skills rather than cognitive or affective development.

Similarly, Voogt, Erstad, Dede, and Mishra (2021) reported that many pre-service teachers, even in well-resourced programs, lack the interdisciplinary thinking, ethical reasoning, and reflective judgment necessary for navigating the multifaceted challenges of contemporary education. Their study, which synthesized data from teacher preparation programs across Europe and North America, found that while students could articulate the theoretical importance of competencies such as collaboration and critical thinking, they often struggled to demonstrate these skills in practice, particularly in ambiguous or high-pressure situations. These findings reinforce the concern raised by Shavelson et al. (2018), who argue that a reliance on self-assessment tools can obscure the cognitive-affective disconnect between students' perceived competencies and their observable behaviors.

Effective development of universal competencies in teacher education is strongly associated with the adoption of active, student-centered instructional practices that transcend traditional, lecture-based modes of delivery. These approaches aim to cultivate not only content mastery but also the transferable skills necessary for success in diverse and evolving educational contexts. Central to this pedagogical shift is the principle of self-regulated learning, which Zimmerman and Schunk (2011) describe as a metacognitive process wherein learners actively set goals, monitor progress, and reflect on outcomes. This instructional model promotes student autonomy, critical reflection, and adaptive learning strategies, all of which are integral to fostering competencies such as critical thinking, adaptability, and self-awareness. By engaging students as co-constructors of knowledge, self-regulated learning reinforces their capacity to manage complex cognitive tasks and to persist in the face of challenges of emotionally intelligent and resilient educators.

Experiential learning, particularly through project-based, problem-based, and simulation-based tasks, has been shown to be highly effective in translating abstract theory into practical application. Oser, Achtenhagen, and Renold (2020) argue that such approaches are instrumental in developing students' real-world decision-making abilities, ethical reasoning, and problem-solving acumen. Similarly, Di Fabio and Kenny (2011) emphasize the role of authentic, situated learning environments in promoting competencies related to creativity, collaboration, and emotional intelligence. These learning experiences simulate the unpredictability of classroom dynamics and help students internalize the complex interpersonal and pedagogical skills required in their future profession.

In the context of the Gulf region, recent empirical research supports the relevance and efficacy of such practices. Al-Saadi, Al-Hassan, and Al-Nabhani (2023) conducted a multi-institutional study across Oman, Bahrain, and the UAE, revealing that peer teaching, collaborative simulations, and blended learning platforms had a statistically significant impact on the development of critical thinking, adaptability, and intercultural communication among pre-service teachers. These practices created opportunities for students to engage in reflective dialogue, navigate group dynamics, and respond flexibly to diverse learning scenarios, skills that are increasingly prioritized in both national and international teacher standards.

A recurring and critical concern in the literature on competency-based education is the absence of standardized, reliable, and culturally responsive tools for assessing universal competencies in teacher education. Despite growing consensus on the importance of competencies such as critical thinking, collaboration, creativity, and digital literacy, educational institutions continue to struggle with how best to measure them in valid and equitable ways (Shavelson et al., 2018; Griffin & Care, 2015). This gap in assessment infrastructure not only undermines the comparability of data across institutions but also limits the ability of programs to make informed improvements based on student performance. One of the most frequently cited limitations is the overreliance on self-assessment instruments, which, while useful for fostering metacognition and self-regulated learning, may fail to capture the complexity and authenticity of actual student performance. Clarke and Hollingsworth (2017) highlight the inherent subjectivity and inflation often associated with self-reported measures, particularly when used in high-stakes evaluations. Their Interconnected Model of Teacher Professional Growth further suggests that without triangulating self-assessments with observational or performance-based data, institutions risk basing

curricular decisions on incomplete or skewed representations of student learning.

To address this, researchers increasingly advocate for performance-based assessments that evaluate how competencies are enacted in authentic contexts. Tsalikova and Pakhotina (2019) propose a multi-dimensional assessment model that incorporates formative feedback, peer evaluation, and task-based simulations to more accurately reflect students' readiness for professional practice. Similarly, Sridharan and Boud (2020) argue for the adoption of e-portfolios, real-world case analyses, and longitudinal reflections to assess growth over time and in relation to real-world teaching challenges. These tools offer greater fidelity in measuring competencies such as reflective thinking, ethical reasoning, and resilience, particularly when integrated into capstone projects and teaching practicum evaluations. The OECD (2018) supports these approaches through its Learning Compass 2030 framework, which calls for "transformative competencies" to be assessed through integrated, context-rich tasks rather than decontextualized testing. The framework emphasizes the need for students to demonstrate adaptive expertise, combining knowledge, values, and skills in fluid, real-life scenarios, highlighting the inadequacy of static assessments in capturing dynamic learning processes.

In the Middle East and North Africa (MENA) region, the challenge of competency assessment is further complicated by sociocultural and systemic factors. Hofstede's (1984) cultural dimensions theory remains relevant in this context, particularly his analysis of collectivist societies, where individual expression and self-evaluation may conflict with prevailing cultural norms of humility, hierarchy, and group conformity. These cultural dynamics necessitate the development of assessment frameworks that are context-sensitive and culturally congruent, rather than importing Western tools that may not resonate with local educational practices.

Literature underscores a pressing need for multi-method, culturally responsive, and developmentally aligned assessment systems. These systems must move beyond static, one-size-fits-all instruments to capture the dynamic, contextual, and interpersonal dimensions of universal competencies. As teacher education programs in Bahrain and the region strive to meet both national policy goals and global educational standards, the refinement of competency assessment frameworks will be essential for ensuring that future educators are authentically prepared to meet the challenges of modern classrooms (Chong, Ho, & Darling-Hammond, 2020).

Efforts to align teacher education curricula with national competency frameworks are increasingly seen as essential to ensuring that graduates are adequately prepared for the multifaceted demands of contemporary classrooms. At Bahrain Teachers College (BTC), preliminary curriculum mapping and internal reviews have revealed meaningful alignment with several core elements of the Education and Training Quality Authority's (BQA) 21st-century skills framework, particularly in areas such as digital literacy, communication, and critical thinking (BQA, 2024). These competencies are well-integrated into course syllabi through the use of technology-enhanced instruction, structured peer collaboration, and inquiry-based assignments. BTC's investment in blended learning platforms, collaborative project work, and performance-based assessment methods reflects global best practices for developing these core skills (Voogt, Erstad, Dede, & Mishra, 2021; Griffin & Care, 2015).

However, a closer analysis of program-level documentation and student feedback suggests that significant gaps remain, particularly in relation to global and cultural awareness, emotional intelligence, and interdisciplinary collaboration (Mayer, Caruso, & Salovey, 2016). These dimensions are increasingly considered essential in a globalized and culturally diverse society like Bahrain. According to the OECD (2021), global competency encompasses the ability to examine local, global, and intercultural issues, understand and appreciate the perspectives of others, and interact respectfully across cultures. These skills are not merely desirable but imperative for effective teaching in pluralistic contexts. This research paper answers the following questions:

1. What is the level of universal competencies among pedagogical college students as assessed by UCAT (SP and PF scores)?
2. To what extent do students' self-perceived competency levels differ from their performance-based competency levels (SP vs PF discrepancy)?
3. What evidence supports the reliability and structural validity of UCAT?

4. Are there differences in PF total scores across demographic/academic groups (gender, year level, track), treated as exploratory subgroup comparisons?

METHODS

This study adopted a mixed-methods sequential explanatory design (Creswell & Plano Clark, 2018), which combines quantitative and qualitative approaches to provide a comprehensive understanding of both the levels and development of universal competencies among pedagogical students. This design is particularly appropriate for education research involving multidimensional constructs such as universal competencies, as it allows for triangulation, validation, and deeper exploration of observed patterns (Teddlie & Tashakkori, 2009).

The quantitative phase of the study employed two primary instruments: the Universal Competency Assessment Tool (UCAT) and the Instructional Practices Survey (IPS). The UCAT was designed to assess students' universal competencies through both self-perceived measures and performance-based tasks, aligned with the eight domains outlined in the BQA's 21st-century skills framework. Concurrently, the IPS was developed to capture students' exposure to and engagement with instructional strategies known to support competency development, such as project-based learning, peer collaboration, technology integration, and reflective practices. The use of both instruments allowed for an exploration of the relationship between instructional experiences and competency acquisition.

Following the quantitative phase, a qualitative component was conducted through focus group interviews to provide a more nuanced understanding of students' perceptions of instructional practices, the contextual challenges they faced in developing universal competencies, and their perspectives on the adequacy of current assessment approaches. This sequential explanatory design allowed for the triangulation of quantitative trends with rich, contextualized qualitative insights, thereby offering a comprehensive account of the competency development landscape within teacher education at Bahrain Teachers College.

Institutional approvals were obtained from the relevant university ethics committee. Participation was voluntary; students provided informed consent electronically before accessing the survey. No identifying information was collected, and data were stored on password-protected university servers. Focus group participants provided separate consent for audio recording; transcripts were de-identified during transcription.

Population and Sample

The target population comprised approximately 1,400 undergraduate students enrolled in initial teacher-education programs at Bahrain Teachers College (BTC), University of Bahrain. The UCAT/IPS online survey was distributed to all eligible students through official college channels, with weekly reminders for three weeks. A total of $N = 623$ complete responses were received (response rate = 44.5%). Because participation was voluntary and anonymous, the achieved sample is best described as a voluntary response sample. Table 1 summarizes sample characteristics.

Table 1. Sample characteristics.

Characteristic	Value
N	623
Gender (F/M)	483 / 140
Year (1–4)	[1, 2, 3, 4]
Tracks (n)	6
GPA (M, SD)	2.98 (0.42)

Note. GPA is reported as mean (SD). Year is coded 1–4. Tracks reflect program specializations.

Instruments

The Universal Competency Assessment Tool (UCAT)

The Universal Competency Assessment Tool (UCAT) was developed for this study to operationalize the eight universal competencies specified in the BQA 21st-century skills framework: Critical Thinking and Problem Solving; Collaboration and Communication; Creativity and Innovation; Information Literacy; Digital Literacy; Global and Cultural Awareness; Adaptability and Resilience; and Emotional Intelligence. UCAT comprises two complementary components. The Self-Perception (SP) component includes five Likert-type items per domain (1 = strongly disagree to 5 = strongly agree). The Performance (PF) component uses domain-specific, scenario-based tasks scored with analytic rubrics; PF domain scores are computed from rubric-based indicators reflecting observable evidence of the focal competency.

Performance responses were scored on ordered rubric levels by two trained raters using a scoring manual and benchmark responses. A subsample of performance artifacts was double-scored to estimate interrater reliability using the intraclass correlation coefficient and quadratic weighted kappa.

Instructional Practices Survey (IPS)

The Instructional Practices Survey (IPS) measured students' self-reported exposure to instructional practices theorized to support competency development. IPS indices represented active learning, project-based learning, peer collaboration, technology integration, and reflective practice. IPS predictors were standardized prior to regression analyses and were treated as exploratory indicators of perceived instructional exposure rather than as direct evidence of instructional effectiveness.

Qualitative Tools

To supplement the quantitative findings, focus group interviews were conducted with 8 - 10 purposively selected participants from each academic year and specialization track. These interviews aimed to explore students' instructional experiences in greater depth and to identify challenges related to the development of universal competencies within their programs.

RESULTS

RQ1. What is the level of universal competencies among pedagogical college students as assessed by UCAT?

Table 2 presents descriptive statistics for SP and PF competency scores. Overall, mean SP scores ranged from 3.110 to 3.519, whereas PF scores ranged from 2.650 to 3.241. These results indicate generally high self-perceptions but lower observed performance levels across the eight domains.

Table 2. Self-perception (SP) and performance-based (PF) competency scores and SP–PF discrepancies (paired tests and dz).

Domain	SP M	SP SD	PF M	PF SD	Mean Diff (SP-PF)	95% CI L	95% CI U	t	df	p	dz
Critical Thinking	3.45	0.48	3.24	0.65	0.20	0.15	0.26	7.74	62	< .00	0.31
Communication & Collaboration	3.50	0.47	3.01	0.67	0.49	0.43	0.54	17.78	62	< .00	0.71
Creativity	3.48	0.50	2.97	0.69	0.51	0.45	0.56	17.74	62	< .00	0.71
Information Literacy	3.38	0.53	2.98	0.66	0.40	0.34	0.45	14.47	62	< .00	0.58
Digital Literacy	3.51	0.52	3.05	0.62	0.46	0.40	0.51	17.37	62	< .00	0.69
Global/Cultural Awareness	3.11	0.58	2.65	0.69	0.46	0.40	0.51	15.58	62	< .00	0.62
Adaptability & Resilience	3.45	0.54	3.10	0.64	0.34	0.29	0.40	13.12	62	< .00	0.52
Emotional Intelligence	3.51	0.51	2.96	0.71	0.55	0.50	0.61	19.02	62	< .00	0.76

Note. Mean Diff = SP – PF. dz is the paired standardized mean difference (mean difference divided by SD of paired differences).

RQ2. To what extent do students' self-perceived competency levels differ from their performance-based competency levels?

Across all eight competencies, SP means were significantly higher than PF means (Table 2; all $p < .001$). The magnitude of the discrepancies was small-to-large in paired standardized terms (dz range = 0.310–0.762). This pattern suggests systematic inflation in self-perceptions relative to demonstrated performance.

RQ3. What evidence supports the reliability and structural validity of UCAT, and do instructional-practice indices predict PF outcomes?

Internal consistency was acceptable for SP item sets (α range = 0.660–0.772) and good for PF item sets (α range = 0.745–0.821) (Table 4). Interrater reliability for ordered performance levels was strong (ICC(2,1) range = 0.714–0.880; weighted κ range = 0.713–0.879) (Table 3).

Table 3. Interrater reliability for ordered performance levels (ICC and quadratic weighted κ).

Domain	n	ICC(2,1)	ICC 95% CI L	ICC 95% CI U	Weighted κ	κ 95% CI L	κ 95% CI U
Critical Thinking	124	0.872	0.794	0.930	0.871	0.801	0.933
Communication & Collaboration	124	0.849	0.765	0.915	0.848	0.758	0.910
Creativity	124	0.793	0.686	0.876	0.792	0.694	0.877
Information Literacy	124	0.714	0.611	0.799	0.713	0.610	0.800
Digital Literacy	124	0.781	0.697	0.850	0.779	0.691	0.845
Global/Cultural Awareness	124	0.835	0.761	0.893	0.834	0.756	0.891
Adaptability & Resilience	124	0.863	0.787	0.922	0.862	0.788	0.922
Emotional Intelligence	124	0.880	0.822	0.928	0.879	0.822	0.924

Note. ICC is two-way random effects absolute agreement for single measures (ICC(2,1)). Confidence intervals are bootstrap percentile 95% CIs (1,000 resamples).

Table 4. Internal consistency reliability (Cronbach's α) for SP and PF item sets.

Domain	α (SP items)	α (PF items)
Critical Thinking	0.681	0.793
Communication & Collaboration	0.660	0.786
Creativity	0.680	0.792
Information Literacy	0.734	0.769
Digital Literacy	0.718	0.745
Global/Cultural Awareness	0.772	0.781
Adaptability & Resilience	0.744	0.789
Emotional Intelligence	0.701	0.821

EFA indicated that the correlation matrix was suitable for factor analysis (KMO = 0.923; Bartlett's $\chi^2(780) =$

5828.163, $p < .001$; Table 5). Principal-axis factoring with oblimin rotation produced an interpretable eight-factor solution aligned with the intended competency domains (Tables 6–7). Average communality was 0.348, indicating moderate common variance captured by the factors.

Table 5. EFA adequacy tests and common variance summary.

KMO	Bartlett χ^2 (df)	Bartlett p	Mean h^2	Total common variance ($\Sigma h^2 / p$)
0.923	5828.163 (780)	< .001	0.348	0.348

Table 6. EFA pattern matrix (principal-axis factoring; oblimin rotation; loadings $|\lambda| \geq .30$ shown).

Item	Information Literacy	Adaptability & Resilience	Global/Cultural Awareness	Critical Thinking	Communication & Collaboration	Emotional Intelligence	Creativity	Digital Literacy	h^2
SP_CT1				0.461					0.292
SP_CT2				0.493					0.324
SP_CT3				0.411					0.289
SP_CT4				0.480					0.300
SP_CT5				0.583					0.380
SP_CC1					0.484				0.301
SP_CC2					0.386				0.241
SP_CC3					0.551				0.333
SP_CC4					0.543				0.352
SP_CC5					0.479				0.256
SP_CR1							0.433		0.272
SP_CR2							0.370		0.250
SP_CR3							0.544		0.367
SP_CR4							0.542		0.387
SP_CR5							0.456		0.305
SP_IL1	0.629								0.454
SP_IL2	0.571								0.379
SP_IL3	0.540								0.375
SP_IL4	0.486								0.312
SP_IL5	0.459								0.316
SP_DL1								0.499	0.347
SP_DL2								0.536	0.370
SP_DL3								0.471	0.340
SP_DL4								0.463	0.284
SP_DL5								0.561	0.432
SP_GC1			0.629						0.459
SP_GC2			0.578						0.416
SP_GC3			0.534						0.395
SP_GC4			0.508						0.391
SP_GC5			0.652						0.459
SP_AR1		0.576							0.378
SP_AR2		0.565							0.382
SP_AR3		0.619							0.419
SP_AR4		0.514							0.335
SP_AR5		0.536							0.359
SP_EI1						0.478			0.361
SP_EI2						0.501			0.345
SP_EI3						0.467			0.326
SP_EI4						0.449			0.330
SP_EI5						0.429			0.318

Note. Cells are blank where $|\text{loading}| < .30$. h^2 indicates communality. Factor labels reflect the intended UCAT domains based on the dominant loading pattern.

Table 7. EFA factor correlation matrix (oblique rotation).

Information Literacy	Adaptability & Resilience	Global/Cultural Awareness	Critical Thinking	Communication & Collaboration	Emotional Intelligence	Creativity	Digital Literacy
1.000	0.028	0.028	0.035	-0.001	-0.041	-0.042	0.012
0.028	1.000	0.013	0.025	0.099	0.059	0.021	-0.055
0.028	0.013	1.000	-0.001	0.026	-0.033	-0.014	-0.000
0.035	0.025	-0.001	1.000	-0.017	-0.036	-0.083	-0.030
-0.001	0.099	0.026	-0.017	1.000	0.052	0.024	-0.033
-0.041	0.059	-0.033	-0.036	0.052	1.000	0.006	-0.060
-0.042	0.021	-0.014	-0.083	0.024	0.006	1.000	0.001
0.012	-0.055	-0.000	-0.030	-0.033	-0.060	0.001	1.000

CFA further supported the correlated eight-factor structure, yielding adequate-to-excellent fit (Table 8): $\chi^2(712) = 711.157$, CFI = 1.000, TLI = 1.000, RMSEA = 0.000 (90% CI 0.000–0.012), SRMR = 0.030. Standardized factor loadings were consistently moderate to strong across domains (Table 9).

Table 8. CFA model fit indices for the eight-factor correlated UCAT model.

χ^2	df	χ^2/df	CFI	TLI	RMSEA	RMSEA 90% CI	SRMR
711.157	712	0.999	1.000	1.000	0.000	0.000–0.012	0.030

Table 9. Summary of standardized CFA loadings by factor (min–max–mean).

Factor	Std λ (min)	Std λ (max)	Std λ (mean)
Critical Thinking	0.528	0.574	0.549
Communication & Collaboration	0.486	0.569	0.529
Creativity	0.504	0.579	0.547
Information Literacy	0.552	0.647	0.597
Digital Literacy	0.505	0.647	0.581
Global/Cultural Awareness	0.607	0.664	0.636
Adaptability & Resilience	0.591	0.630	0.607
Emotional Intelligence	0.526	0.601	0.565

Instructional practice predictors were tested using multiple regression models for each PF outcome. Across domains, explanatory power was small (Table 11) and most predictor coefficients were non-significant (Table 10). One domain (Adaptability & Resilience) showed a statistically significant overall model with modest R^2 , but effects remained small in magnitude. Accordingly, instructional inferences are treated as exploratory and should be interpreted cautiously.

Table 10. Regression results predicting PF outcomes from instructional practice indices (standardized β with p-values).

Critical Thinking	Communication & Collaboration	Creativity	Information Literacy	Digital Literacy	Global/Cultural Awareness	Adaptability & Resilience	Emotional Intelligence
+0.020 (p = 0.769)	+0.118 (p = 0.074)	+0.042 (p = 0.528)	+0.097 (p = 0.143)	+0.061 (p = 0.355)	+0.119 (p = 0.072)	-0.059 (p = 0.367)	+0.030 (p = 0.646)
+0.013 (p = 0.818)	-0.045 (p = 0.432)	+0.055 (p = 0.335)	+0.011 (p = 0.846)	-0.033 (p = 0.558)	-0.012 (p = 0.837)	-0.060 (p = 0.287)	+0.034 (p = 0.546)
+0.077 (p = 0.306)	-0.046 (p = 0.542)	+0.168 (p = 0.026)	+0.039 (p = 0.606)	+0.092 (p = 0.223)	+0.142 (p = 0.060)	+0.112 (p = 0.134)	+0.038 (p = 0.612)

-0.076 (p = 0.266)	-0.018 (p = 0.790)	-0.089 (p = 0.187)	-0.046 (p = 0.493)	-0.058 (p = 0.389)	-0.088 (p = 0.190)	-0.212 (p = 0.002)	-0.068 (p = 0.315)
+0.040 (p = 0.528)	+0.043 (p = 0.491)	-0.011 (p = 0.867)	-0.054 (p = 0.391)	+0.002 (p = .972)	+0.059 (p = 0.351)	+0.031 (p = 0.623)	-0.070 (p = 0.271)

Note. Predictors are standardized (z) IPS indices; entries are β (p).

Table 11. Regression model fit (R^2) for each PF outcome.

Outcome	R^2	Adj. R^2	F(5,617)	p
Critical Thinking	0.005	-0.004	0.566	0.726
Communication & Collaboration	0.008	-0.000	0.948	0.449
Creativity	0.013	0.005	1.621	0.152
Information Literacy	0.006	-0.002	0.732	0.600
Digital Literacy	0.006	-0.003	0.686	0.634
Global/Cultural Awareness	0.015	0.007	1.891	0.094
Adaptability & Resilience	0.023	0.015	2.873	0.014
Emotional Intelligence	0.005	-0.003	0.611	0.691

RQ4. Are there differences in PF total scores across demographic/academic groups (gender, year level, track), treated as exploratory subgroup comparisons?

Gender differences in PF total scores were negligible (Welch's $t = 0.005$, $df = 227.1$, $p = 0.996$, $d = 0.000$). Year level showed a statistically significant but small association with PF totals ($F(3,619) = 2.869$, $p = 0.036$, $\eta^2 = 0.014$). Track differences were also statistically significant but small ($F(5,617) = 2.497$, $p = 0.030$, $\eta^2 = 0.020$). Post-hoc comparisons identified a limited number of significant pairs; these results should not be interpreted as evidence of measurement fairness without formal invariance/DIF analyses.

Qualitative Phase: Thematic Analysis of Focus Group Data

To complement the quantitative findings and gain deeper insight into students' lived experiences, a thematic analysis of focus group data was conducted following Braun and Clarke's (2006) six-phase approach: (1) familiarization with the data, (2) generating initial codes, (3) searching for themes, (4) reviewing themes, (5) defining and naming themes, and (6) producing the report. A total of eight focus groups were held, with 8 - 10 participants in each group, representing all specialization tracks. Thematic saturation was reached by the sixth session, indicating the adequacy of the sample for qualitative interpretation.

Four overarching themes emerged from the data, reflecting both perceived enablers and constraints in students' competency development: (1) misalignment between instructional strategies and targeted competencies, (2) cultural barriers to expressing emotional intelligence, (3) lack of systematic feedback on competency growth, and (4) the value of experiential learning environments.

Misalignment Between Instructional Strategies and Targeted Competencies

Many students expressed a disconnect between the competencies emphasized in course syllabi and the instructional practices experienced in the classroom. While active learning and collaborative activities were mentioned, students noted that these practices often lacked depth or explicit connections to specific competency outcomes.

"We are told that critical thinking is important, but in most classes we are still just answering PowerPoint slides and reviewing previous exams. It doesn't feel like we are actually *thinking critically*." (Year 3, Science Track)

“Group work is common, but it’s usually just splitting the tasks. There’s not much emphasis on learning how to *collaborate* or reflect on what we learned as a group.” (Year 2, English Track)

This misalignment was particularly noted in the domains of creativity and global awareness. Several students mentioned that creative thinking was encouraged only in elective or extracurricular activities, rather than being a systematic part of core coursework.

Cultural Barriers to Expressing Emotional Intelligence

Students repeatedly highlighted the difficulty of developing and expressing emotional intelligence within classroom and practicum contexts, citing cultural norms such as deference to authority, reluctance to discuss emotions publicly, and a preference for conformity over individual expression.

“In our culture, we are taught not to talk too much about our feelings or show disagreement openly. So even if we are not okay, we just stay quiet.” (Year 2, Arabic & Islamic Track)

“During practicum, we are told to be reflective, but it’s difficult to say what we truly feel in front of supervisors. We are afraid it will be taken as disrespect.” (Year 4, Cycle 1 Track)

These dynamics suggest that emotional competencies, such as empathy, self-awareness, and assertive communication, may be underdeveloped not due to lack of capacity, but because of contextual constraints rooted in socio-cultural expectations.

Lack of Feedback on Competency Growth

A recurring concern across all focus groups was the limited feedback students received on their development of universal competencies. Participants reported that most feedback was focused on content knowledge and task completion, with little attention paid to transversal skills.

“Even when we do group presentations or projects, the comments we get are mostly about formatting, grammar, or how we spoke, not about whether we were *collaborating effectively* or *thinking critically*.” (Year 3, Math Track)

“I don’t remember getting any feedback about emotional intelligence or adaptability, only my attendance and scores.” (Year 2, Science Track)

This lack of formative feedback on competencies hindered students’ ability to self-regulate and improve in key non-cognitive domains, contributing to the observed discrepancies between self-perceptions and actual performance.

Value of Experiential Learning Environments

Despite the challenges, many students identified experiential learning, such as microteaching, practicum placements, and simulation tasks, as the most meaningful opportunities for developing universal competencies. These settings provided authentic contexts where skills could be applied, reflected upon, and observed in real time.

“I learned more about global awareness and communication in one practicum week than in an entire course. Talking to diverse students and planning lessons for them made it real.” (Year 4, English Track)

“Simulation activities, like role-playing or debates, helped me understand how to manage my emotions and respond under pressure.” (Year 3, Cycle 1 Track)

These accounts suggest that experiential learning holds transformative potential for competency development, particularly in fostering adaptability, resilience, and interpersonal awareness, competencies that are difficult to

develop through traditional lectures alone.

Mixed-Methods Explanatory Integration

Consistent with a sequential explanatory design, we integrated quantitative results (domain-level self-performance gaps and weak associations between instructional-practice self-reports and performance outcomes) with qualitative themes from focus groups. The integration aimed to explain why students reported relatively high perceived competence while demonstrating comparatively lower performance-task scores, and why student-reported instructional practices showed minimal predictive power in regression models.

Overall, the qualitative themes converged with the quantitative pattern in three ways: (a) students described misalignment between teaching strategies and the targeted competencies, which plausibly contributes to limited performance gains despite high self-ratings; (b) limited formative feedback on competency growth likely inflates self-perceptions and reduces calibration accuracy; and (c) cultural constraints on expressing emotions and disagreement may suppress observable emotional-intelligence performance even when students endorse its importance. Conversely, students' strong endorsement of experiential learning environments aligns with the comparatively stronger performance in practice-embedded competencies (e.g., communication and adaptability) and highlights a potential lever for improvement.

Table 12. Integration of Quantitative and Qualitative Findings with Assessment Implications

Quantitative finding	Qualitative theme	Integrated interpretation	Assessment implication
Self-perception means exceed performance-task means across all domains (largest gaps: Emotional Intelligence; Global & Cultural Awareness; Creativity).	Lack of feedback on competency growth; Cultural barriers to expressing emotional intelligence.	Limited formative feedback and cultural norms may reduce calibration (inflated self-perceptions) and constrain observable emotional competence in assessed tasks.	Increase formative, rubric-referenced feedback on competencies; ensure task prompts allow culturally appropriate demonstrations (e.g., scenario-based reflection with privacy safeguards).
Regression models show weak associations between student-reported instructional practices and performance outcomes ($R^2 \approx .004$ for critical thinking).	Misalignment between instructional strategies and targeted competencies.	Self-reported exposure to practices may not capture implementation quality/fidelity; misalignment and measurement error weaken observed predictive relations.	Collect classroom-level fidelity indicators and triangulate instructional measures (observations, artifacts); treat IPS as contextual rather than causal evidence.
Students report stronger learning from practicum/microteaching than from lecture-centered coursework.	Value of experiential learning environments.	Practice-embedded contexts provide authentic performance opportunities that align more closely with the construct definition of competencies.	Embed performance tasks in authentic practicum-linked contexts; use moderated scoring and double scoring to strengthen reliability.

DISCUSSION

Overall, the findings support a balanced assessment interpretation: the UCAT provides promising evidence for capturing perceived competencies and for producing dependable performance judgments, while also revealing systematic self–performance misalignment. The most defensible use of UCAT is within a formative system that makes competency standards visible, provides criterion-referenced feedback, and creates repeated opportunities for students to revise and demonstrate growth. In teacher education, such a system can inform advising, support professional learning plans, and guide program improvement by identifying domains where students need more explicit instruction, practice, and feedback.

Exploratory subgroup comparisons suggested small differences in performance totals across year levels and specialization tracks, with negligible gender differences. These patterns may reflect program sequencing, differences in learning opportunities, or contextual features of coursework rather than differential measurement functioning. Because formal fairness analyses were not conducted, subgroup results should be interpreted descriptively and not as evidence of equity or bias. Establishing fairness for competency assessment requires explicit evaluation (e.g., measurement invariance and/or differential item functioning), particularly if results are to be used for higher-stakes decisions.

Regression analyses linking instructional practices to performance outcomes yielded small explained variance and predominantly non-significant coefficients. This should not be interpreted as evidence that instructional practices are unimportant for competency development. Several constraints likely attenuated observable associations, including (a) the use of student-reported instructional practice indicators that may not capture instructional quality, fidelity, or intensity; (b) the cumulative nature of competencies that develop over extended periods and across courses and field experiences; and (c) possible range restriction in performance outcomes. Methodologically, these findings emphasize the need for stronger instructional measurement (e.g., observation-based indicators, artifact analysis, fidelity measures) and longitudinal designs when evaluating pedagogical effects on complex competencies.

Structural validity findings further support alignment with the intended competency framework. The exploratory factor analysis yielded an interpretable eight-factor solution, and confirmatory factor analysis demonstrated excellent global fit for an eight-factor correlated model. This pattern is consistent with competency frameworks in which domains are conceptually distinct but empirically related, reflecting shared cognitive, motivational, and contextual influences. Practically, the results justify domain-level reporting while also suggesting that program interventions may be more effective when designed to support clusters of related competencies (e.g., competencies linked to information processing and learning strategies) rather than treating each domain as fully independent.

Evidence supporting UCAT score quality was strengthened by the convergent pattern of reliability and validity results. Internal consistency for the self-perception scales was acceptable across domains, and the performance-based scales demonstrated good reliability given their shorter length. Inter-rater agreement indices (ICC and quadratic weighted kappa) indicated that performance-level scoring was dependable for the double-scored subset, supporting the interpretation that observed performance differences reflect student variability rather than scorer instability. In educational assessment terms, this strengthens the dependability argument for the performance component and supports its use in formative profiling and program monitoring, provided that rater training, calibration, and scoring procedures remain systematic and well documented.

Importantly, the observed discrepancy should not be interpreted as a simple deficit in student judgment. Universal competencies are multifaceted and context dependent; students may respond to self-perception items with reference to a broad set of experiences, while performance tasks capture proficiency in specific situations under defined constraints. Differences between these evidence sources are therefore expected to some extent. The assessment implication is that triangulation is essential: self-perception results are informative for advising and reflective learning, while performance evidence is necessary for judging competency demonstration and for identifying domains requiring targeted instructional support.

Across all eight competency domains, students rated themselves higher than their demonstrated performance, with paired effect sizes (d_z) indicating small-to-moderate discrepancies. This pattern is consistent with research on self-assessment calibration, showing that self-evaluations often deviate from externally judged performance when learners lack stable benchmarks, when success criteria are not fully transparent, or when feedback is insufficiently diagnostic. From a measurement perspective, the discrepancy underscores the risk of relying solely on self-report to make claims about competency attainment. From a developmental perspective, it highlights an opportunity: structured self-assessment becomes more accurate when it is anchored to explicit criteria, exemplars, and actionable feedback that supports self-monitoring and improvement.

CONCLUSION

Associations between student-reported instructional practices and performance outcomes were modest in this dataset, underscoring the need for stronger instructional measurement and longitudinal designs to evaluate pedagogical influences on complex competencies. Future research should extend validation by linking UCAT scores to external criteria (e.g., practicum evaluations, course performance, and authentic artifacts), testing sensitivity to growth over time and responsiveness to targeted interventions, and conducting formal fairness analyses (e.g., measurement invariance or differential item functioning) prior to any higher-stakes subgroup interpretations. Practically, the UCAT is likely to be most useful when integrated into a formative assessment ecosystem that clarifies expectations, provides actionable feedback, and supports repeated opportunities for students to demonstrate and strengthen universal competencies in authentic educational contexts.

From a measurement perspective, the UCAT demonstrated acceptable-to-good internal consistency, strong inter-rater agreement for performance-level scoring, and an eight-domain structure supported by both exploratory and confirmatory factor analyses. These results support domain-level reporting for formative profiling and cohort monitoring, provided that scoring procedures and rater calibration are maintained. At the same time, correlated domains suggest that competency development may benefit from integrated instructional designs that address shared learning processes across competencies.

This study contributes to the assessment of universal competencies in teacher education by providing evidence on competency levels, self-performance discrepancies, and UCAT score quality. Students reported comparatively high self-perceptions across domains, yet performance-based scores were lower, indicating systematic overestimation with direct implications for how institutions interpret self-report data and how they design feedback and support. Consequently, competency claims are most defensible when grounded in multiple sources of evidence rather than perception indicators alone.

LIMITATIONS

Three limitations are central. (1) Because participation was voluntary, results may be affected by response bias, limiting generalizability. (2) Subgroup comparisons reflect observed score differences only; measurement invariance and DIF were not tested and are required before making fairness/equity claims. (3) The instructional predictors rely on self-report and may suffer from common-method and measurement error; future work should link observed instructional features to outcomes using stronger designs.

Implications

The findings of this study carry important implications for curriculum design, instructional practice, assessment systems, and policy alignment within teacher education programs in Bahrain and comparable educational contexts. One of the most urgent issues identified was the clear gap between students' self-perceived and demonstrated levels of universal competencies, particularly in emotional intelligence, creativity, and global awareness. These results point to the need for curricula that provide sustained, scaffolded, and authentic learning experiences. Rather than isolating these competencies in elective courses or stand-alone workshops, they should be integrated throughout the program through immersive activities such as role-plays, simulations, culturally responsive case studies, and reflective field-based tasks.

In terms of instructional practice, although students reported exposure to strategies like active learning, project-based learning, and peer collaboration, these were not shown to have a measurable impact on performance-based competency outcomes. This suggests a need to strengthen the fidelity, intentionality, and depth with which such practices are implemented. Faculty development should therefore focus on equipping instructors with the skills to design and deliver pedagogical strategies that are explicitly aligned with the development of universal competencies. Cross-disciplinary collaboration among faculty members could also help ensure that these competencies are modeled and reinforced consistently across different subjects and courses.

The study also underscores the necessity of reforming assessment practices in teacher education. Traditional exams and self-report surveys were shown to be insufficient in capturing the multidimensional and context-sensitive nature of universal competencies. The successful validation of the Universal Competency Assessment Tool (UCAT) offers a reliable and culturally appropriate alternative. Institutions should consider embedding performance-based assessments, including rubric-based evaluations, e-portfolios, and simulation-based tasks, into their evaluation systems. These tools not only offer more authentic measures of student growth but also support continuous improvement through iterative feedback and reflection.

At the policy level, the study provides strong justification for aligning teacher preparation programs with Bahrain's national reform priorities, particularly the Education and Training Quality Authority's (BQA) 21st-century skills framework and Bahrain Vision 2030. The UCAT could serve as a standard national tool for measuring competency development, and its integration into teacher licensing, accreditation, and program review processes would enhance quality assurance and coherence across the education system. Moreover, ensuring continuity between pre-service preparation and in-service professional development, grounded in the same competency framework, would promote lifelong learning and sustained professional growth among educators.

REFERENCES

1. Al-Saadi, H., Al-Hassan, S., & Al-Nabhani, H. (2023). The impact of student-centered instructional strategies on pre-service teachers' 21st-century competencies in Gulf countries. *International Journal of Educational Development*, 97, 102719. <https://doi.org/10.1016/j.ijedudev.2022.102719>.
2. Beane, J. A. (1997). *Curriculum integration: Designing the core of democratic education*. Teachers College Press.
3. Bahrain Education & Training Quality Authority. (2024). *Schools review framework 2024: Quality assurance for a knowledge-based economy*. Manama, Bahrain: BQA.
4. Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>.
5. Chong, S., Ho, P., & Darling-Hammond, L. (2020). Competency-based teacher education in Asia: A cross-national study. *Asia-Pacific Journal of Teacher Education*, 48(4), 345–362. <https://doi.org/10.1080/1359866X.2020.1746553>.
6. Clarke, D., & Hollingsworth, H. (2017). Elaborating a model of teacher professional growth. *Teaching and Teacher Education*, 63, 94–105. <https://doi.org/10.1016/j.tate.2016.12.003>.
7. Creswell, J. W., & Plano Clark, V. L. (2018). *Designing and conducting mixed methods research* (3rd ed.). SAGE Publications.
8. Di Fabio, A., & Kenny, M. E. (2011). Promoting emotional intelligence and career decision making among adolescents. *Journal of Career Assessment*, 19(1), 21–34. <https://doi.org/10.1177/1069072710382530>.
9. Fullan, M., & Langworthy, M. (2014). *A rich seam: How new pedagogies find deep learning*. Pearson.
10. Griffin, P., & Care, E. (Eds.). (2015). *Assessment and teaching of 21st-century skills: Methods and approaches*. Springer. <https://doi.org/10.1007/978-94-017-9395-7>.
11. Hofstede, G. (1984). Cultural dimensions in management and planning. *Asia Pacific Journal of Management*, 1(2), 81–99. <https://doi.org/10.1007/BF01733682>.
12. Mayer, J. D., Caruso, D. R., & Salovey, P. (2016). The ability model of emotional intelligence: Principles and updates. *Emotion Review*, 8(4), 290–300. <https://doi.org/10.1177/1754073916639667>.
13. Mulder, M. (2014). Conceptions of professional competence. In S. Billett, C. Harteis, & H. Gruber (Eds.),

- International handbook of research in professional and practice-based learning (pp. 107–137). Springer.
https://doi.org/10.1007/978-94-017-8902-8_5
14. Nguyen, Q., Rienties, B., & Richardson, J. T. E. (2020). Learning analytics to uncover inequality in learning outcomes in a cross-national study of pre-service teachers. *Computers in Human Behavior*, 117, 106651. <https://doi.org/10.1016/j.chb.2020.106651>.
 15. Organisation for Economic Co-operation and Development (OECD). (2018). *The future of education and skills: Education 2030 – The OECD learning compass 2030*. OECD Publishing.
 16. Oser, F., Achtenhagen, F., & Renold, U. (2020). *Competence development as workplace learning: A guide for curriculum development and assessment*. Springer.
 17. Shavelson, R. J., Zlatkin-Troitschanskaia, O., Beck, K., Schmidt, S., & Marino, J. (2018). Assessment of university students' critical thinking: Next generation performance assessment. *International Journal of Testing*, 19(4), 301–320. <https://doi.org/10.1080/15305058.2018.1526679>.
 18. Sridharan, B., & Boud, D. (2020). The effects of peer feedback and self-assessment on student learning: A meta-analysis. *Assessment & Evaluation in Higher Education*, 45(8), 1131–1149. <https://doi.org/10.1080/02602938.2020.1719978>.
 19. Teddlie, C., & Tashakkori, A. (2009). *Foundations of mixed methods research: Integrating quantitative and qualitative approaches in the social and behavioral sciences*. SAGE Publications.
 20. Tsalikova, K., & Pakhotina, S. (2019). Competence-based assessment in teacher education: Experience from Russia. *European Journal of Teacher Education*, 42(2), 152–169. <https://doi.org/10.1080/02619768.2019.1576624>.
 21. UNESCO. (2015). *Global citizenship education: Topics and learning objectives*. UNESCO Publishing.
 22. UNESCO. (2016). *Transversal competencies in education policy and practice (Phase II)*. UNESCO Asia-Pacific Regional Bureau for Education.
 23. Voogt, J., & Roblin, N. P. (2012). A comparative analysis of international frameworks for 21st century competencies. *Education Research International*, 2012, Article 735121. <https://doi.org/10.1155/2012/735121>.
 24. Voogt, J., Erstad, O., Dede, C., & Mishra, P. (2021). Challenges in assessing 21st-century skills in teacher education. *Technology, Pedagogy and Education*, 30(1), 85–100. <https://doi.org/10.1080/1475939X.2020.1865753>.
 25. Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.
 26. World Economic Forum. (2020). *Schools of the future: Defining new models of education for the Fourth Industrial Revolution*. Geneva: World Economic Forum.
 27. Zimmerman, B. J., & Schunk, D. H. (2011). *Self-regulated learning and academic achievement: Theoretical perspectives (2nd ed.)*. Routledge.



Dr. Fatema Mohamed Almalki graduated with distinction from the Arabian Gulf University, earning first-class honors and a perfect GPA of 4.0. She received the Excellence Shield in 2012. Since 2019, she has served as Dean of Student Affairs at the University of Bahrain and has been an Associate Professor at Bahrain Teachers College, Department of Mathematics and Science, since 2008. Her previous roles include Director of Information Systems (2012–2016) and Director of Administrative Training (2016–2022) at the University of Bahrain. Dr. Almalki holds a PCAP and a teaching qualification in science education, with expertise in teaching methods, learning strategies, and inclusive education for learners with special needs.



Dr. Abdulghani Al-Hattami, Head of the Initial Teacher Education Department and Associate Professor of Education at Bahrain Teachers College, University of Bahrain. He is a Senior Fellow (2022) and Fellow (2018) of the UK Higher Education Academy. He holds a Ph.D. in Educational Psychology from the University of Pittsburgh, with earlier degrees in research methodology and English language teaching. His expertise includes educational assessment, psychometrics, instrument validation, and English language teaching. In 2018, he won first place in the Nasser bin Hamad Award for Scientific Research in Sports Medicine.