

Adaptation and validation of HowULearn 2 student questionnaire for use in Italian tertiary education

Shaniko Kaleci · Stefania Costi · Pierantonio Bellini · Luigi Generali · Antonella Lotti · Paola Sena · Johanna Chester · Giovanna Amaducci · Daniela Mecugni

Received: 14 October 2025 / Accepted: 29 September 2026

© The Author(s) 2026

Abstract

Background Healthcare students face multiple academic, clinical, and psychosocial stressors that may affect their well-being and academic performance. The *HowULearn 2* questionnaire, developed at the University of Helsinki, is a research-based tool designed to explore students' experiences of teaching and learning and to provide feedback that supports self-regulated learning and teaching improvement. However, the questionnaire has not yet been validated for the Italian language and cultural context. This study aimed to translate, culturally adapt, and evaluate the *HowULearn 2* questionnaire for Italian undergraduate healthcare students, assessing face and content validity, structural validity, internal consistency, and test–retest reliability.

Methods Face validity was assessed in 40 students and content validity by an eight-member expert panel. Structural validity was examined using exploratory factor analysis with principal-axis factoring and oblique Promax rotation. Internal consistency was assessed using Cronbach's alpha, and test–retest reliability using intraclass correlation coefficients (ICC [3, 1]) after four weeks.

Results Of 351 invited students, 345 provided questionnaire data; 343 (99.42%) had ≤ 6 missing item responses, and 294 complete cases were included in the exploratory factor analysis. Face validity was supported: 52 of 61 items (85.2%) met the predefined clarity criterion. Content validity was high (S-CVI/Ave = 0.91). Test–retest reliability was assessed in 63 participants who completed both administrations. Horn's parallel analysis, based on 1,000 random replications and the 95th-percentile criterion, supported seven factors. A five-factor solution was retained as a theory-driven exploratory solution consistent with the intended questionnaire dimensions. KMO was 0.923 and Bartlett's test was statistically significant ($\chi^2 = 11,533.725$, $df = 1,830$, $p < 0.001$). Cronbach's alpha coefficients ranged from 0.785 to 0.930 across subscales (overall alpha = 0.952). Test–retest ICCs ranged from 0.607 to 0.838.

Conclusions The Italian *HowULearn 2* showed satisfactory face and content validity, high sampling adequacy, and a theoretically interpretable five-factor exploratory solution, with acceptable-to-excellent internal consistency and moderate-to-good test–retest reliability. Further independent and multicentre studies, including confirmatory factor analysis, are warranted to assess the stability and generalisability of the measurement structure.

Keywords Validation · Questionnaire · Learning · Students · Higher education · Quality assurance · Structural validity

This Article in Press is shared early to give you faster access to new research. It is citable and carries a permanent DOI. The final edited version will replace it automatically.



Introduction

Healthcare students are exposed to numerous academic, clinical, and psychosocial stressors that can negatively affect their physical and mental well-being, as well as their academic performance and satisfaction with learning environments [1]. Student academic delay and dropout are key indicators of potential shortcomings in the teaching–learning process, often reflecting gaps in institutional support and instructional quality, and academic progression [2, 3]. Therefore, systematic monitoring of progress and experiences is essential for ensuring high-quality higher education.

Quality assurance (QA) in higher education refers to systematic processes and policies designed to monitor, evaluate, and improve teaching, learning, and institutional performance. QA frameworks help universities identify weaknesses, support evidence-based decision-making, and promote student success. International bodies such as UNESCO, the OECD, the European Union, and the European Association for Quality Assurance in Higher Education (ENQA) have emphasized the importance of QA in fostering continuous improvement, accountability, and alignment with global standards [4–6].

The Bologna Process, initiated in 1999 by 29 European countries, aimed to harmonize higher education systems within the European Higher Education Area (EHEA), introducing a three-cycle degree structure (bachelor's, master's, doctorate), the European Credit Transfer and Accumulation System (ECTS), and shared quality assurance standards. This process also promoted a student-centered learning culture, emphasizing learning outcomes, feedback, and reflective practice as essential components of teaching quality. Research has shown that student-centered approaches and effective feedback mechanisms are associated with greater engagement, satisfaction, and academic achievement [7–9].

In Italy, universities have progressively implemented QA frameworks and teaching and learning centers to enhance instructional quality. However, the student perspective, particularly regarding their learning processes, self-efficacy, and perceptions of teaching—remains underrepresented in institutional evaluations. Instruments that provide structured, research-based insights into students' learning experiences are therefore essential to improving educational quality and reducing attrition.

The HowULearn questionnaire, developed by the University of Helsinki, was designed to explore the interrelationships between students' approaches to learning, perceptions of teaching, academic burnout, self-efficacy, and metacognitive awareness [10, 11]. The updated version, *HowULearn 2*, includes 61 items across five main domains: (1) Study and Learning I; (2) Study and Learning II; (3) Development of Teaching; (4) Study and Learning III; and (5) Self-assessment of Study, plus an additional section on work-life competence. The questionnaire provides both individual and aggregated feedback to students and teachers, helping identify strengths and areas for improvement. It has been validated and used in several countries, including Finland, Australia, and Singapore, but not yet in Italy [12].

In line with the University of Modena and Reggio Emilia's 2020–2025 Strategic Plan to enhance learning outcomes and reduce dropout rates, this study aimed to adapt and evaluate the *HowULearn 2* questionnaire for use in the Italian higher education context.

Objective

The objective of this study was to translate and culturally adapt the *HowULearn 2* questionnaire into Italian and to evaluate its face and content validity, structural validity, internal consistency, and test–retest reliability among undergraduate healthcare students. The objective of this study was to translate and culturally adapt the *HowULearn 2* questionnaire into Italian and to evaluate its face and content validity, structural validity, internal consistency, and test–retest reliability among undergraduate healthcare students.

Methods

Study design

This methodological study followed the seven-phase framework for cross-cultural adaptation and psychometric evaluation of self-report instruments based on the recommendations of Beaton et al. (2000) and Sousa & Rojanasrirat (2011) [13, 14]. The phases included: (1) Forward translation, (2) Synthesis of translations, (3) Back translation, (4) Expert review and content validation, (5) Pretesting and face validation, (6) Administration of the final questionnaire to the target population, and (7) Assessment of test-retest reliability.

The study protocol was reviewed within the FAR 2021 project evaluation procedure by the Departmental Research Commission of the Dipartimento Chirurgico, Medico, Odontoiatrico e di Scienze Morfologiche con interesse Trapiantologico, Oncologico e di Medicina Rigenerativa (CHIMOMO), University of Modena and Reggio Emilia (UNIMORE), on October 12, 2021.

The study was conducted as an institutional/educational research project within the University Research Fund (FAR) 2021 programme of UNIMORE. The project, entitled “HowULearn: validazione e adattamento cross-culturale di un questionario sugli studenti universitari.” According to the applicable institutional procedures for this type of institutional/educational research, the project did not require review by an Ethics Committee. All participants received information about the study procedures and participation was voluntary. Informed consent was obtained before questionnaire completion. Permission to use and culturally adapt the HowULearn 2 questionnaire was obtained from Anna Parpala (University of Helsinki, Finland) during a dedicated HowULearn project meeting held on 16 December 2021, at which the Italian adaptation and validation project was presented. Permission was also obtained to reproduce the HowULearn questionnaire items included in the current submission.

Study participants

The target population comprised undergraduate students enrolled in five healthcare-related programmes at the Faculty of Medicine and Surgery at the University of Modena and Reggio Emilia (UNIMORE), Italy. Eligible students were enrolled in the first, second, or third year of five programmes: Nursing, Physiotherapy, Medicine and Surgery, Biomedical Laboratory Techniques, and Dental Hygiene.

The planned sample size was 351 students. The final questionnaire was administered to the eligible population through institutional e-mail. A total of 345 students provided questionnaire data and were included in the main psychometric analyses, after application of the predefined data-completeness criteria.

Two subsamples were selected: Face-validation group: 40 students (approximately 6% of total) selected through a random block sampling procedure stratified by program and year of study.

Test-retest group: 80 students (12%) randomly selected with similar stratified procedure.

Inclusion criteria were: being a currently enrolled undergraduate student, Italian native speaker, and consenting to participate. There were no exclusion criteria related to gender or academic performance.

Validation team

The validation team was composed of eight members, ensuring both linguistic and disciplinary expertise: two bilingual translators fluent in Italian and English; two subject-matter experts in pedagogy and educational sciences; two healthcare educators experienced in questionnaire development and validation; one psychometrician; and the principal investigator.

Each member independently reviewed translated versions for semantic, idiomatic, experiential and conceptual equivalence. The team met in three consensus rounds to discuss discrepancies and refine item wording. Items were rated for relevance, clarity, and cultural appropriateness on a four-point scale.

Content validity was quantified using the Item-level Content Validity Index (I-CVI) and scale-level Content Validity Index (S-CVI). Items with I-CVI < 0.78 were revised or reworded until consensus was achieved.

HowULearn 2 questionnaire

The students were instructed to respond to the following 61 items on a Likert scale: completely disagree, disagree, neither agree nor disagree, agree and completely agree. As some items were reversed, the interpretation of the results considered all the responses to be completely negative, negative, neutral, positive or completely positive, independent of the question structure. Reverse-worded items were recoded before calculation of the corresponding subscale scores and before the psychometric analyses, so that higher scores reflected the intended direction of each dimension.

The questionnaire was evaluated according to its five proposed dimensions, and psychometric analyses were primarily conducted at the subscale level. The total questionnaire score was examined descriptively but was not interpreted as representing a single unidimensional construct unless supported by the psychometric findings.

Translation and synthesis process

Phase I: forward translation

Two bilingual translators independently translated the original English questionnaire into Italian, documenting any ambiguities or cultural nuances. One translator had a health-education background, whereas the other was a lay translator. The independent translations were compared and discrepancies, ambiguities, and culturally sensitive expressions were discussed to produce a synthesized Italian version.

Phase II: synthesis of translation

A consensus meeting between translators and researchers produced a synthesized Italian version, incorporating the most accurate and culturally appropriate wording. A written synthesis report documented all decisions and justifications.

Phase III: back translation

The synthesized version was then independently back-translated into English by two translators who were blinded to the original questionnaire. Discrepancies between the back-translated and original versions were reviewed with attention to semantic, idiomatic, experiential, and conceptual equivalence. Cultural and linguistic adaptations were made when required while preserving the intended meaning of the original items.

Expert validation and pretesting

Phase IV: validation team

The validation team reviewed all versions, establishing equivalence at four levels: *Semantic equivalence* (same meaning of words), *Idiomatic equivalence* (replacement of colloquial expressions), *Experiential equivalence* (relevance of items to Italian academic context), and *Conceptual equivalence* (preservation of theoretical meaning).

All items reached consensus for equivalence. The overall S-CVI/Ave was 0.91, indicating excellent content validity.

Phase V: pretesting (face validity)

The pre-final questionnaire was administered to 40 students selected from the target population. Each item was rated as “clear” or “unclear.” The questionnaire was considered acceptable if > 80% of items were rated as clear by at least 60% of participants. Feedback was used to refine minor linguistic elements; no conceptual revisions were required. The resulting version was finalized as the Italian HowULearn 2 questionnaire.

Main administration and reproducibility

Phase VI: administration of the questionnaire to the target population

The final Italian version of the questionnaire was administered electronically using Google Forms to the 351 eligible/invited students through institutional e-mail. A total of 345 students provided questionnaire data. Of these, 343 (99.4%) had six or fewer missing item responses and remained eligible for analyses requiring the predefined completeness criterion, whereas two participants were excluded from such analyses. The EFA was conducted on 294 complete cases using listwise deletion. Questionnaire responses collected were linked to the respondents' student number; student numbers connected to the responses were available to the data manager only (SK), who did not have access to any personal data associated with the student number. All derived data were anonymized before analysis. Participation was voluntary, and informed consent was obtained before questionnaire completion.

Phase VII: test - retest reliability

Eighty students were selected for the test–retest assessment using a stratified random block procedure designed to obtain representation from the participating degree programmes and years of study. Selected students received an e-mail invitation containing a link to complete the questionnaire again.

The second administration took place 4 weeks after the initial assessment. The same 61-item Italian HowULearn 2 questionnaire was administered at both time points.

Of the 80 students invited, 63 completed both administrations and were therefore included in the test–retest analysis, corresponding to a completion rate of 78.8%.

Test–retest reliability was assessed using intraclass correlation coefficients (ICCs) among participants with complete data at both administrations.

The seven phases of the validation process are described in Fig. 1.

Sample size calculation

According to Hatcher and O'Rourke [15], a minimum of five participants per item and at least 100 participants are required for exploratory factor analysis (EFA). Given the 61 items of the questionnaire, a minimum analytic sample of 305 participants was therefore estimated. Assuming an expected non-response rate of 15%, the final target recruitment sample was increased to 351 participants.

Sample adequacy was considered in conjunction with sample size, communalities, factor-loading magnitude, number of factors, number of indicators per factor, and factor overdetermination. The final sample was considered adequate for the exploratory analyses, while recognizing that EFA does not provide definitive confirmation of the factor structure.

The adequacy of the final sample for factor analysis was subsequently evaluated empirically using the Kaiser–Meyer–Olkin (KMO) measure of sampling adequacy, Bartlett's test of sphericity, communalities, and the observed factor-loading pattern.

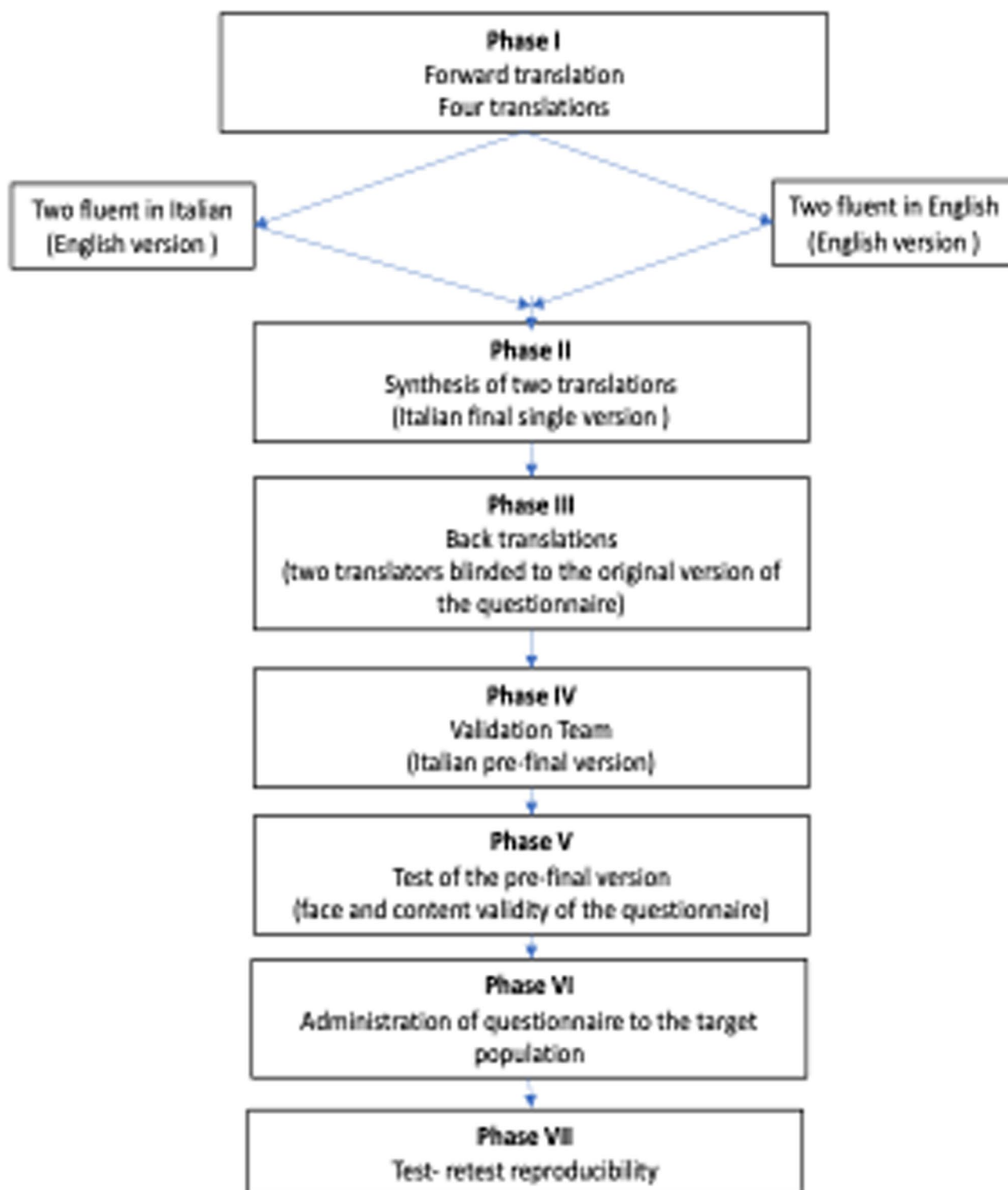


Fig. 1 Seven-phase process for the cross-cultural adaptation and psychometric evaluation of the Italian HowULearn 2 questionnaire

Statistical analysis

Descriptive statistics were calculated for each questionnaire item, including the mean, standard deviation, median, range, and response distribution. Floor and ceiling effects were considered present when $\geq 15\%$ of respondents selected the lowest or highest response category, respectively [16].

Following phase VII, missing responses were assessed. Respondents with more than six missing item responses were considered insufficiently complete for the corresponding scale-level analyses. The distribution of missing responses is reported in Supplementary Table S5; the EFA used 294 complete cases (listwise deletion).

Corrected item–rest correlations were calculated to evaluate the contribution of each item to its corresponding subscale. These correlations were interpreted descriptively, with higher values indicating stronger relationships between the item and the corresponding subscale.

The structural validity of the Italian version was investigated using exploratory factor analysis (EFA). Because EFA is exploratory, the analysis was used to examine the observed dimensional structure in relation to the five-dimensional structure proposed for the original HowULearn 2, rather than to confirm that structure.

Factor extraction was performed using principal axis factoring (PAF). Because the dimensions assessed by the questionnaire may be conceptually related, an oblique rotation (Promax) was used to allow correlations between factors. The adequacy of the data for factor analysis was assessed using the Kaiser–Meyer–Olkin (KMO) measure and Bartlett’s test of sphericity. KMO values ≥ 0.80 were considered indicative of good sampling adequacy, whereas values ≥ 0.90 were considered excellent [17]. Bartlett’s test was considered statistically significant at $p < 0.05$.

The number of factors was evaluated using the conventional eigenvalue criterion, Horn’s parallel analysis, the scree plot, and theoretical interpretability. Horn’s parallel analysis was conducted using 1,000 random replications and the 95th-percentile criterion. In the unrotated principal-factor analysis, the first eight eigenvalues were 18.013, 5.245, 3.234, 1.794, 1.657, 1.329, 1.220, and 1.076; the conventional eigenvalue > 1 criterion would therefore suggest eight factors. The adjusted eigenvalues from Horn’s parallel analysis were 17.4647, 4.7460, 2.8341, 1.4347, 1.3522, 1.0838, 1.0450, and 0.9088 for factors 1–8, respectively, supporting retention of seven factors. Because the original questionnaire specifies five principal dimensions, a five-factor solution was examined as a theory-driven exploratory solution. Factor loadings, communalities, uniquenesses, and cross-loadings were examined. Internal consistency was assessed using Cronbach’s alpha for each of the five subscales. The number of items comprising each subscale was reported together with the corresponding Cronbach’s alpha and 95% confidence interval, where available. Alpha values ≥ 0.70 were considered acceptable, ≥ 0.80 good, and ≥ 0.90 excellent [17]. Very high alpha values were interpreted cautiously because they may partly reflect redundancy among items rather than only high internal consistency.

Because the questionnaire is multidimensional, internal consistency was primarily evaluated at the subscale level. The overall questionnaire alpha was reported descriptively but was not interpreted as evidence of unidimensionality.

Test–retest reliability was assessed using intraclass correlation coefficients (ICCs) based on a two-way mixed-effects model assessing absolute agreement for single measurements [ICC(3,1)]. The analysis included the 63 participants who completed both administrations, conducted four weeks apart. ICCs were calculated separately for each of the five questionnaire dimensions. Paired t-tests were additionally used to assess systematic differences in mean scores between baseline and retest.

ICC values were interpreted according to established reliability guidelines, with values in the 0.50–0.75 range generally indicating moderate reliability and values above 0.75 indicating good reliability [19].

Test–retest reliability was evaluated separately for each subscale. The overall questionnaire score was not used for reliability interpretation because the questionnaire is multidimensional.

Statistical analysis was performed using STATA[®] software version 17 (StataCorp. 2021. Stata Statistical Software: Release 17. College Station, TX: StataCorp LLC.).

Results

A total of 351 students were invited to participate. Of these 345 (98.3%) provided questionnaire data. Two participants had more than six missing responses and were therefore excluded from analyses requiring the predefined completeness criterion; 343 (99.4%) had six or fewer missing responses. The EFA was conducted on 294 complete cases.

Face and content validity

Face validity was assessed during pretesting: 52 out of 61 items (85.2%) met the predefined clarity criterion. Minor grammatical adjustments were made, but no conceptual modifications were required (Table 1).

The validation team calculated an overall Scale-Level Content Validity Index (S-CVI/Ave) of 0.91, indicating a high level of agreement regarding the relevance, clarity, and cultural appropriateness of the questionnaire items (Table 2).

Structural validity

Corrected item–rest correlations were generally positive within the five dimensions, although some items showed weaker associations than others. Item-level corrected item–rest correlations and Cronbach's alpha if the item was deleted are reported in Supplementary Table S3. No item was removed solely on the basis of a single statistical criterion; item retention considered the theoretical content of the questionnaire and the cross-cultural adaptation objective.

The structural validity of the Italian HowULearn 2 questionnaire was examined using exploratory factor analysis (EFA) with principal-axis factoring and oblique Promax rotation. The overall KMO was 0.923 and Bartlett's test of sphericity was statistically significant ($\chi^2 = 11,533.725$, $df = 1,830$, $p < 0.001$), indicating that the correlation matrix was suitable for factor analysis. The analysis included 294 complete cases.

A five-factor solution was examined because it corresponds to the theoretical structure proposed for the original questionnaire. The full Promax pattern matrix, including loadings below $|0.30|$ and cross-loadings, communalities, and uniquenesses, is reported in Supplementary Table S2; the unrotated eigenvalues are reported in Supplementary Table S2a. The five factors showed correlations ranging from -0.245 to 0.537 (Supplementary Table S2b). The scree plot was also inspected as a descriptive aid to factor retention. Because the factor analysis was exploratory and CFA was not performed, the five-factor solution is presented as exploratory evidence rather than confirmation of the hypothesized structure.

The conventional eigenvalue > 1 criterion suggested eight factors. Horn's parallel analysis, based on 1,000 random replications using the 95th-percentile criterion, supported the retention of seven factors. The scree plot was also inspected as a descriptive aid. Because the original questionnaire specifies five principal dimensions, a five-factor solution was examined as a theory-driven exploratory solution.

Reliability and internal consistency

Internal consistency for the questionnaire was assessed using Cronbach's alpha. Cronbach's alpha coefficients ranged from 0.785 to 0.930 across the five subscales: 0.785 for Study and Learning I, 0.873 for Study and Learning II, 0.918 for Development of Teaching, 0.897 for Study and Learning III, and 0.930 for Self-assessment of Study (Table 3). The overall questionnaire alpha was 0.952, but this coefficient was reported descriptively and was not interpreted as evidence of unidimensionality.

Floor and ceiling effects were assessed at the item level using a predefined threshold of $\geq 15\%$ of responses in the lowest or highest response category, respectively. Several items exceeded this threshold; the complete item-level results are reported in Supplementary Table S4.

Table 1 Student responses to the HowULearn 2 academic questions ($n = 345$)

	Test for clarity	Students' responses					Mean±SD (min-max)
	Clear <i>n</i> (%)	Com- pletely disagree <i>n</i> (%)	Disagree <i>n</i> (%)	Neither agree nor disagree <i>n</i> (%)	Agree <i>n</i> (%)	Com- pletely agree <i>n</i> (%)	
Part 1: Studying and learning I							
1. I often have trouble making sense of the things I have to learn.*	26 (65.0)	42 (12.7)	145 (42.0)	81 (23.5)	62 (17.9)	15 (4.3)	2.6±1.0 (1–5)
2. I put a lot of effort into my studying.	38 (95.0)	7 (2.0)	7 (2.0)	31 (9.0)	193 (56.1)	106 (30.8)	4.1±0.8 (1–5)
3. Much of what I've learned seems no more than unrelated bits and pieces.*	20 (50.0)	35 (10.2)	151 (43.9)	79 (23.0)	68 (19.8)	11 (3.2)	2.6±1.0 (1–5)
4. On the whole, I've been systematic and organized in my studying.	38 (95.0)	10 (2.9)	27 (7.8)	81 (23.5)	176 (51.1)	50 (14.5)	3.6±0.9 (1–5)
5. Ideas and perspectives I've come across while I'm studying make me contemplate them from all sides.	28 (70.0)	4 (1.2)	31 (9.0)	123 (35.9)	151 (44.1)	33 (9.6)	3.5±0.8 (1–5)
6. I look at evidence carefully to reach my own conclusion about what I'm studying.	35 (87.5)	7 (2.0)	16 (4.7)	63 (18.4)	205 (59.8)	52 (15.1)	3.8±0.8 (1–5)
7. I am unable to understand the topics I need to learn because they are so complicated.*	19 (47.5)	65 (18.8)	168 (48.7)	80 (23.2)	29 (8.4)	3 (0.9)	2.2±0.8(1–5)
8. I organize my study time carefully to make the best use of it.	35 (87.5)	11 (3.2)	41 (11.9)	89 (25.8)	148 (42.9)	56 (16.2)	3.5±1.0 (1–5)
9. Often I have to repeat things in order to learn them.	35 (87.5)	8 (2.3)	19 (5.5)	44 (12.8)	175 (50.9)	98 (28.5)	3.9±0.9 (1–5)
10. I carefully prioritize my time to make sure I can fit everything in.	36 (90.0)	18 (5.2)	61 (17.7)	64 (18.6)	147 (42.7)	54 (15.7)	3.4±1.1(1–5)
11. I try to relate new material to my previous knowledge.	37 (92.5)	8 (2.3)	8 (2.3)	46 (13.4)	203 (59.2)	78 (22.7)	3.9±0.8(1–5)
12. I try to relate what I have learned in one course to what I learn in other courses.	37 (92.5)	8 (2.3)	10 (2.9)	47 (13.6)	201 (58.3)	76 (22.9)	3.9±0.8(1–5)
Part 2: Studying and learning II							
1. I believe I will do well in my studies	34 (85.0)	8 (2.3)	14 (4.0)	86 (24.9)	191 (55.4)	46 (13.3)	3.7±0.8(1–5)
2. I'm certain I can understand the most difficult material in my studies	33 (82.5)	5 (1.4)	21 (6.1)	112 (32.6)	170 (49.4)	36 (10.5)	3.6±0.8(1–5)
3. I'm confident I can understand the basic concepts of my own study field	38 (95.0)	8 (2.3)	6 (1.7)	15 (4.4)	219 (63.8)	95 (27.7)	4.1±0.7(1–5)
4. I expect to do well in my studies	35 (87.5)	8 (2.3)	5 (1.5)	40 (11.7)	205 (59.9)	84 (24.6)	4.0±0.7(1–5)
5. I'm certain I can learn well the skills required in my study field	39 (97.5)	9 (2.6)	13 (3.8)	46 (13.3)	202 (58.5)	75 (21.7)	3.9±0.8(1–5)
Part 3: Development of teaching							
Arrangement and structure							
1. It is clear to me what I am expected to learn in the courses	36 (90.0)	11 (3.2)	30 (8.7)	75 (21.8)	183 (53.2)	45 (13.0)	3.6±0.9(1–5)
2. What we are taught seems to match what we are supposed to learn.	30 (75.0)	17 (4.9)	31 (8.9)	94 (27.2)	162 (46.9)	41 (11.9)	3.5±0.9(1–5)
Teaching and learning							
3. I can see the relevance of what we are taught	32 (80.0)	9 (2.6)	23 (6.7)	98 (28.6)	172 (50.1)	41 (11.9)	3.6±0.8(1–5)
4. Students support each other and try to give help when it is needed.	36 (90.0)	13 (3.8)	29 (8.4)	66 (19.1)	174 (50.4)	63 (18.3)	3.7±0.9(1–5)
5. I find most of what I learned in courses really interesting.	38 (95.0)	12 (3.5)	15 (4.4)	66 (19.2)	185 (53.9)	65 (18.9)	3.8±0.9(1–5)
6. Talking with other students helps me to develop my understanding.	40 (100)	5 (1.4)	12 (3.5)	47 (13.6)	181 (52.5)	100 (28.9)	4.0±0.8(1–5)
7. I enjoy participating in courses	34 (85.0)	5 (1.5)	17 (4.9)	79 (23.0)	181 (52.8)	61 (17.8)	3.8±0.8(1–5)
8. I can generally work comfortably with other students	39 (97.5)	11 (3.2)	18 (5.2)	55 (16.0)	176 (51.3)	83 (24.2)	3.8±0.8(1–5)
Requirements and Feedback							

Table 1 (continued)

	Test for clarity	Students' responses					Mean ± SD (min-max)
	Clear <i>n</i> (%)	Com- pletely disagree <i>n</i> (%)	Disagree <i>n</i> (%)	Neither agree nor disagree <i>n</i> (%)	Agree <i>n</i> (%)	Com- pletely agree <i>n</i> (%)	
9. I receive enough feedback about my learning.	27 (67.5)	18 (5.2)	55 (15.9)	141 (41.0)	113 (32.8)	17 (4.9)	3.1±0.9(1–5)
10. It is clear to me what is expected in the assessed work (i.e. final exam. exercises).	33 (82.5)	11 (3.2)	35 (10.2)	88 (25.7)	178 (52.0)	30 (8.8)	3.5±0.8(1–5)
11. I can see how the set exercises fit in with what we are supposed to learn.	38 (95.0)	11 (3.2)	25 (7.3)	83 (24.2)	192 (55.9)	32 (9.3)	3.6±0.8(1–5)
12. The feedback given helps me to improve my ways of learning and studying.	34 (85.0)	9 (2.6)	29 (8.5)	95 (27.9)	170 (49.8)	38 (11.1)	3.5±0.8(1–5)
13. The set work (ex. Exams, exercises, course performances) helps me to make connections to my existing knowledge or experience.	39 (97.5)	10 (2.9)	14 (4.1)	82 (23.9)	197 (57.4)	40 (11.7)	3.7±0.8(1–5)
14. The feedback given on my set course exercises helps me to clarify things I hadn't fully understood	34 (85.0)	14 (4.1)	22 (6.4)	110 (32.2)	165 (48.2)	31 (9.1)	3.5±0.8(1–5)
15. The learning objectives have been clearly specified.	33 (82.5)	9 (2.6)	24 (6.9)	77 (22.3)	193 (55.9)	42 (12.2)	3.7±0.8(1–5)
16. The teaching strategies assist me in achieving the learning objectives of the modules.	33 (82.5)	8 (2.3)	28 (8.2)	99 (28.9)	178 (52.0)	29 (8.5)	3.5±0.8(1–5)
17. The assessment focuses on skills based on the learning objectives of the modules.	33 (82.5)	13 (3.8)	18 (5.2)	83 (24.2)	196 (57.1)	33 (9.6)	3.6±0.8(1–5)
Part 4: Studying and learning III							
1. I feel overwhelmed by the work related to my studies.*	30 (75.0)	10 (2.9)	67 (19.5)	121 (35.3)	97 (28.3)	48 (13.9)	3.3±1.0(1–5)
2. I feel a lack of study motivation and often think of giving up.*	17 (42.5)	78 (22.6)	137 (39.7)	66 (19.1)	46 (13.3)	18 (5.2)	2.3±1.1(1–5)
3. I often have feelings of inadequacy in my studies.*	24 (60)	57 (16.6)	109 (31.8)	66 (19.2)	84 (24.5)	27 (7.9)	2.7±1.2(1–5)
4. I often sleep badly because of matters related to my studies.*	24 (60)	65 (18.9)	97 (28.2)	75 (21.8)	80 (23.3)	27 (7.8)	2.7±1.2(1–5)
5. I feel that I am losing interest in my studies.*	17 (42.5)	114 (33.2)	135 (39.4)	56 (16.3)	30 (8.7)	8 (2.3)	2.0±1.0(1–5)
6. I'm continually wondering whether my studies have any meaning.*	17 (42.5)	112 (32.5)	114 (33.0)	60 (17.4)	41 (11.9)	18 (5.2)	2.2±1.1(1–5)
7. I brood over matters related to my studies during my free time.*	26 (65.0)	59 (17.1)	98 (28.4)	62 (17.9)	95 (27.5)	31 (9.0)	2.8±1.2(1–5)
8. I used to have higher expectations of my studies than I do now.*	21 (52.5)	70 (30.7)	106 (30.7)	50 (14.5)	78 (22.6)	41 (11.9)	2.7±1.3(1–5)
9. The pressure of my studies causes me problems in my close relationships with others.*	22 (55.0)	86 (24.9)	103 (29.9)	68 (19.7)	53 (15.4)	35 (10.1)	2.5±1.2(1–5)
Part 5: Self-Assessment of studying							
1. I can motivate myself to learn when I need to.	34 (85.0)	4 (1.2)	20 (5.8)	61 (17.7)	208 (60.5)	51 (14.8)	3.8±0.7(1–5)
2. I use my intellectual strengths to compensate for my weaknesses.	38 (95.0)	6 (1.7)	18 (5.2)	56 (16.3)	208 (60.5)	56 (16.3)	3.8±0.7(1–5)
3. I know when each strategy I use will be most effective.	32 (80.0)	5 (1.4)	31 (9.0)	98 (28.5)	167 (48.5)	43 (12.5)	3.6±0.8(1–5)
4. I am good at organizing information.	34 (85.0)	5 (1.5)	20 (5.8)	87 (25.4)	195 (57.0)	35 (10.2)	3.7±0.7(1–5)
5. I am good at remembering information.	31 (77.5)	11 (3.2)	30 (8.7)	92 (26.7)	174 (50.4)	38 (11.0)	3.5±0.9(1–5)
6. I am a good judge of how well I understand something.	38 (95.0)	6 (1.7)	10 (2.9)	59 (17.1)	215 (62.5)	54 (15.7)	3.8±0.7(1–5)
7. I try to use strategies that have worked in the past.	40 (100)	5 (1.5)	7 (2.0)	22 (6.4)	237 (69.3)	71 (20.7)	4.0±0.6(1–5)
8. I am aware of what strategies I use when study.	40 (100)	4 (1.2)	15 (4.4)	47 (13.7)	217 (63.4)	59 (17.2)	3.9±0.7(1–5)
9. I find myself using helpful learning strategies automatically.	33 (82.5)	8 (2.3)	24 (6.9)	85 (24.7)	181 (52.6)	46 (13.4)	3.6±0.8(1–5)

Table 1 (continued)

	Test for clarity	Students' responses					
	Clear <i>n</i> (%)	Com- pletely disagree <i>n</i> (%)	Disagree <i>n</i> (%)	Neither agree nor disagree <i>n</i> (%)	Agree <i>n</i> (%)	Com- pletely agree <i>n</i> (%)	Mean $\pm$ SD (min-max)
10. I think about what I really need to learn before I begin a task.	35 (87.5)	9 (2.6)	44 (12.8)	87 (25.4)	164 (47.8)	39 (11.4)	3.5 $\pm$ 0.9(1–5)
11. I set specific goals before I begin a task.	33 (82.5)	12 (3.5)	43 (12.6)	70 (20.5)	176 (51.5)	41 (11.9)	3.5 $\pm$ 0.9(1–5)
12. I organize my time to best accomplish my goals.	39 (97.5)	9 (2.6)	32 (9.3)	65 (18.9)	181 (52.6)	57 (16.6)	3.7 $\pm$ 0.9(1–5)
13. I change strategies when I fail to understand.	39 (97.5)	5 (1.4)	26 (7.6)	81 (23.7)	191 (55.8)	39 (11.4)	3.6 $\pm$ 0.8(1–5)
14. I ask myself how what I'm reading is related to what I already know.	39 (97.5)	3 (0.9)	8 (2.3)	69 (20.0)	212 (61.6)	52 (15.1)	3.8 $\pm$ 0.7(1–5)
15. I find myself analyzing the usefulness of strategies while I study.	32 (80.0)	12 (3.5)	28 (8.2)	90 (26.2)	181 (52.8)	32 (9.3)	3.5 $\pm$ 0.9(1–5)
16. I summarize what I've learned after I finish.	35 (87.5)	13 (3.8)	42 (12.2)	46 (13.4)	170 (49.4)	73 (21.2)	3.7 $\pm$ 1.0(1–5)
17. I ask myself how well I accomplish my goals.	36 (90.0)	10 (2.9)	17 (5.0)	66 (19.3)	195 (57.2)	53 (15.5)	3.7 $\pm$ 0.8(1–5)
18. I ask myself if I learned as much as I could have once I finish a task.	39 (97.5)	7 (2.0)	20 (5.8)	59 (17.2)	198 (57.9)	58 (16.9)	3.8 $\pm$ 0.8(1–5)

* Reverse-worded items. Responses are reported on a five-point Likert scale. Face-validity ratings were obtained from the 40-student pretest-ing group

Test–retest reliability

Test–retest reliability was assessed four weeks after the initial administration. Eighty students were invited to participate in the second administration, and 63 participants completed both administrations and were included in the test–retest analysis. The completion rate was 78.8% (63/80).

Test–retest reliability was assessed using intraclass correlation coefficients (ICCs) based on a two-way mixed-effects model with absolute agreement for single measurements [ICC(3,1)]. ICCs ranged from 0.607 to 0.838 across the five subscales. The ICCs were 0.679 (95% CI 0.522–0.792) for Study and Learning I, 0.614 (95% CI 0.436–0.747) for Study and Learning II, 0.644 (95% CI 0.472–0.768) for Development of Teaching, 0.838 (95% CI 0.746–0.899) for Study and Learning III, and 0.607 (95% CI 0.424–0.742) for Self-assessment of Study (Table 3).

No statistically significant differences in mean scores were observed between baseline and retest for any of the five subscales: Study and Learning I (mean difference = -0.050 , $p=0.153$), Study and Learning II (-0.086 , $p=0.129$), Development of Teaching (-0.016 , $p=0.724$), Study and Learning III (0.026 , $p=0.658$), and Self-assessment of Study (-0.034 , $p=0.487$).

Questionnaire responses

Item-level response distributions and floor/ceiling results are reported in the supplementary Tables S1 and S4. Several items exceeded the predefined $\geq 15\%$ threshold for the lowest or highest response category, particularly in Study and Learning III and Self-assessment of Study (Supplementary Table S4).

The overall HowULearn 2 mean score was 3.6 ± 0.4 , corresponding to a response between “neither agree nor disagree” and “agree” (data not shown). The total questionnaire score was not interpreted as a single overall construct because the questionnaire is multidimensional.

Subscale mean item scores in the main questionnaire sample ranged from 2.62 to 3.87 on the 1–5 response scale. Mean item scores were 3.45 ± 0.46 for Study and Learning I, 3.87 ± 0.67 for Study and Learning II, 3.63 ± 0.60 for Development of Teaching, 2.62 ± 0.88 for Study and Learning III, and 3.72 ± 0.58 for Self-assessment of Study. These descriptive results are reported separately from the test–retest baseline means shown in Table 3. These

Table 2 b. Student responses to the HowULearn 2 employment-related questions ($n = 345$)

PART 6: Working during studies		<i>n</i> (%)
1. Have you worked during your studies?	No	186 (53.9)
	Yes	159 (46.1)
2. Please estimate how many hours per week you have worked in the last year.	None	8 (5.0)
	Less than 20 h	53 (33.1)
	20 to 35 h	64 (40.0)
	More than 35 h	35 (21.9)
3. My current job is related to my studies	Completely disagree	83 (52.2)
	I neither agree nor disagree	26 (16.3)
	Completely agree	50 (31.4)
4. I can utilise things I have learnt at the University in my current job.	Completely disagree	74 (47.1)
	I neither agree nor disagree	28 (17.8)
	Completely agree	55 (35.0)
5. The requirement level of my current job corresponds to my university education.	Completely disagree	91 (58.3)
	I neither agree nor disagree	32 (20.5)
	Completely agree	33 (21.2)
6. To which group does your employer belong?	Private company	71 (44.7)
	State public company	22 (13.8)
	Local or regional public company (e.g. municipality, hospital)	13 (8.2)
	Third sector (social cooperative, foundation. etc.)	12 (7.5)
	Freelance	13 (8.2)
	Other	31 (19.5)
7. If you answered “other”, please specify		-
8. What kind of an employment relationship or appointment to a post is your job?	Fixed-term employment relationship	65 (40.9)
	Permanent employment relationship	38 (23.9)
	Other	53 (33.3)
9. If you answered “other”, please specify		-
10. Which of the following best describes your duties in your current job?	Legislators, entrepreneurs and senior management	1 (0.6)
	Intellectual, scientific and highly specialized professions	8 (5.0)
	Technical professions	18 (11.3)
	Executive professions in office work	12 (7.5)
	Qualified professions in commercial activities and services	49 (30.8)
	Artisans, specialized workers and farmers	8 (5.0)
	Plant operators, fixed and mobile machinery workers and vehicle drivers	2 (1.3)
	Unskilled professions	14 (8.8)
	Armed forces	0 (0.0)
	Other	46 (28.9)
11. If you answered “other”; please specify		

Percentages for employment-related questions are calculated among respondents to the relevant item

descriptive results are presented without inferring levels of engagement, metacognitive awareness, or other outcomes that were not directly assessed.

These descriptive findings are consistent with the multidimensional nature of the HowULearn 2 questionnaire and characterize the response patterns observed in this sample.

Table 3 Mean scores, standard deviations, internal consistency, and test–retest reliability of the HowULearn 2 subscales

HowULearn 2 subscale	Items	Mean±SD, baseline	Cronbach's α	Test–retest N	Retest mean±SD	Mean difference†	ICC(3,1) (95% CI)
Study and Learning I	12	3.442±0.301	0.785	63	3.492±0.387	−0.050	0.679 (0.522–0.792)
Study and Learning II	5	3.959±0.487	0.873	63	4.044±0.526	−0.086	0.614 (0.436–0.747)
Development of Teaching	17	3.683±0.427	0.918	63	3.699±0.410	−0.016	0.644 (0.472–0.768)
Study and Learning III	9	2.439±0.815	0.897	63	2.413±0.835	0.026	0.838 (0.746–0.899)
Self-assessment of Study	18	3.780±0.398	0.930	63	3.814±0.457	−0.034	0.607 (0.424–0.742)

† Mean difference calculated as baseline (T0) – retest (T1). ICC(3,1)=two-way mixed-effects model, absolute agreement, single measurement. Test–retest interval: four weeks. No statistically significant baseline–retest differences were observed (all $p>0.12$). Scores are mean item scores on the 1–5 response scale

Discussion

This study provides preliminary evidence regarding the measurement properties of the Italian HowULearn 2 questionnaire. The findings indicate satisfactory internal consistency across the five subscales and moderate-to-good test–retest reliability. The exploratory factor analysis provided evidence of a multidimensional structure. However, both the Kaiser criterion and Horn's parallel analysis indicated a larger empirical dimensionality than the hypothesized five-factor structure. The five-factor solution was therefore retained as a theory-driven exploratory solution rather than as a statistically confirmed model. These findings provide preliminary evidence regarding the measurement properties of the HowULearn 2 questionnaire for assessing students' learning experiences and perceptions of teaching in Italian higher education.

The completion rate was high in both the main administration (98.3%) and the test–retest assessment (78.8%). The cross-cultural adaptation followed Beaton's and Sousa & Rojjanasrirat's guidelines, ensuring conceptual and linguistic equivalence. The expert panel's *S-CVI/Ave of 0.91* provided evidence of high content validity that items were relevant, clear, and culturally appropriate. The expert review supported the relevance, clarity, and cultural appropriateness of the adapted items. Face validity results, with 85% of items rated as “clear” by students, further support the appropriateness and comprehensibility of the Italian version. The exploratory factor analysis provided evidence broadly consistent with the proposed multidimensional structure of HowULearn 2. However, because the present analysis was exploratory, these findings should not be interpreted as confirmatory evidence of the original five-factor model.

Although both the Kaiser criterion and Horn's parallel analysis suggested a larger number of factors, the five-factor solution was retained because it corresponded to the theoretical framework of the original HowULearn 2 questionnaire and yielded the most interpretable structure for cross-cultural adaptation purposes. Consequently, the five-factor model should be regarded as a theory-driven exploratory solution requiring confirmation through CFA in an independent sample.

Internal consistency was satisfactory to high across the five subscales, with Cronbach's alpha coefficients ranging from 0.785 to 0.930. The overall alpha was 0.952 and was reported descriptively because the questionnaire is multidimensional [21, 22]. Similarly, test–retest ICCs ranged from 0.607 to 0.838, indicating moderate-to-good temporal reliability across the five subscales. These findings provide preliminary evidence of internal consistency and temporal reliability, while further validation is needed in independent samples.

The mean overall score was 3.6 ± 0.4 . Because the questionnaire is multidimensional, interpretation focused primarily on the individual subscales. Mean item scores ranged from 2.62 for Study and Learning III to 3.87 for Study and Learning II. These descriptive findings should not be interpreted as direct measures of engagement, metacognitive awareness, satisfaction, or educational burden. These findings align with previous studies demonstrating that students in healthcare fields experience elevated stress and academic burden compared to peers in other disciplines, potentially affecting satisfaction and performance [20]. These descriptive findings should be interpreted as characteristics of the sample rather than as evidence of educational or clinical outcomes.

The present findings provide preliminary exploratory evidence regarding the dimensional structure and measurement properties of the Italian *HowULearn 2* questionnaire.

Furthermore, the lower mean score observed for Study and Learning III should be interpreted descriptively and should not be considered a direct measure of student satisfaction or educational burden. This finding is broadly comparable with observations reported by Parpala et al. (2019).

The Italian version of *HowULearn 2* may be useful for research on students' learning experiences and perceptions of teaching in Italian healthcare education. However, its use in educational quality-improvement or quality-assurance activities should be evaluated in future longitudinal and multicentre studies. The present study did not directly assess changes in teaching quality, student engagement, metacognitive development, or educational outcomes.

The questionnaire may be useful in future educational research and quality-improvement initiatives, but such applications require evaluation in longitudinal and multicentre studies.

These functions are consistent with the *European Standards and Guidelines for Quality Assurance in Higher Education (ESG 2015)* and align with international calls from UNESCO and the OECD to strengthen evidence-based educational policies.

The questionnaire was originally designed to provide information relevant to students' learning approaches and perceptions of teaching; however, the present study did not directly evaluate changes in metacognitive development or educational outcomes. In this sense, the tool contributes to building a culture of student-centered learning, one of the pillars of the Bologna Process.

When used alongside *HowUTeach*—a companion tool for teacher self-assessment—it allows institutions to explore correlations between student and teacher perspectives, promoting a shared vision of teaching and learning quality.

Limitations and future directions

This study has several limitations. First, data were collected at a single Italian university and within healthcare degree programmes, which limits generalizability to other institutions and disciplines. Second, the factor analysis was exploratory and CFA was not conducted; therefore, the hypothesized five-factor structure requires confirmation in an independent sample. Third, the EFA was based on 294 complete cases, which is smaller than the total number of questionnaire respondents. Fourth, item-level floor and ceiling effects were present for several items. Finally, test–retest reliability was moderate for four of the five subscales, and the study was not designed to assess predictive or longitudinal educational outcomes.

Conclusions

The Italian version of the *HowULearn 2* questionnaire provides encouraging preliminary evidence regarding face and content validity, internal consistency, and test–retest reliability in Italian undergraduate healthcare students. The five-factor Promax solution was evaluated as an exploratory, theory-driven model and should not be considered confirmed without CFA. Further validation in independent and multicentre samples is warranted to assess the stability and generalizability of the measurement structure.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1186/s12909-026-10521-2>.

Acknowledgements Not applicable.

Authors' contributions SK: Conceptualization; methodology; validation; formal analysis; investigation; data curation; writing—original draft preparation; writing—review and editing. SC: Conceptualization; investigation; methodology; review

and editing.PB: Conceptualization; investigation; methodology; review and editing.LG: Conceptualization; methodology; review and editing.AL: Conceptualization; investigation; methodology; review and editing.PS: Conceptualization; methodology; review and editing.JC: Review and editing.GA: Conceptualization; investigation; methodology; review and editing.DM: Conceptualization; investigation; methodology; review and editing.All authors read and approved the final manuscript.

Funding This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Data availability The data that support the findings of this study are available from the corresponding author upon reasonable request.

Declarations

Ethics approval and consent to participate The study was conducted in accordance with institutional ethical standards and the 1964 Helsinki Declaration and its later amendments. The study protocol was reviewed by the Departmental Research Commission of the Dipartimento Chirurgico, Medico, Odontoiatrico e di Scienze Morfologiche con interesse Trapiantologico, Oncologico e di Medicina Rigenerativa (CHIMOMO), University of Modena and Reggio Emilia (UNIMORE), on 12 October 2021, as part of the FAR 2021 project evaluation procedure. According to applicable institutional procedures, no Ethics Committee review was required. All participating students provided informed consent prior to participation.

Consent for publication Not Applicable.

Competing interests The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

References

1. Kunzler AM, Helmreich I, König J, Chmitorz A, Wessa M, Binder H, Lieb K. Psychological interventions to foster resilience in healthcare students. *Cochrane Database Syst Rev*. 2020;7(7):CD013684. <https://doi.org/10.1002/14651858.CD013684>. PMID: 32691879; PMCID: PMC7388680.
2. Ribeiro IJS, Pereira R, Freire IV, de Oliveira BG, Casotti CA, Boery EN. Stress and quality of life among university students: A systematic literature review. *Health Prof Educ*. 2018;4:70–7. <https://doi.org/10.1016/j.hpe.2017.03.002>.
3. Simonelli-Muñoz AJ, Balanza S, Rivera-Caravaca JM, Vera-Catala'n T, Lorente AM, Gallego-Go'mez JI. Reliability and validity of the student stress inventory-stress manifestations questionnaire and its association with personal and academic factors in university students. *Nurse Educ Today*. 2018; 64: 156–160. <https://doi.org/10.1016/j.nedt.2018.02.019> PMID: 29494939.
4. UNESCO. Quality assurance and accreditation in higher education: A global perspective. Paris: UNESCO; 2019.
5. European Association for Quality Assurance in Higher Education (ENQA). Standards and Guidelines for Quality Assurance in the European Higher Education Area (ESG 2015). Brussels: ENQA; 2015.
6. OECD. Benchmarking Higher Education System Performance. Paris: OECD Publishing; 2019.
7. European Commission. The Bologna Process and the European Higher Education Area. Available at: <https://www.ehea.info>
8. Biggs J, Tang C. Teaching for Quality Learning at University. 4th ed. McGraw-Hill Education; 2011.
9. Trigwell K, Prosser M. Improving the quality of student learning: The influence of learning context and student approaches to learning. *High Educ*. 1991;22(3):251–66.
10. Parpala A, Lindblom-Ylänne S. Using a research instrument for developing quality at the university. *Qual High Educ*. 2012;18(3):313–28.

11. Ruohoniemi M, Forni M, Mikkonen J, Parpala A. Enhancing quality with a research-based student feedback instrument: a comparison of veterinary students' learning experiences in two European universities. *Qual High Educ.* 2017;23(3):249–63.
12. Parpala A, et al. HowULearn 2: Development and validation of a multidimensional feedback tool. University of Helsinki; 2019.
13. Beaton DE, Bombardier C, Guillemin F, Ferraz MB. (2000). Guidelines for the Process of Cross-Cultural Adaptation of Self-Report Measures. *Spine*, 25(24), 3186–3191, December 15, 2000. <https://doi.org/10.1097/00007632-200012150-00014>
14. Sousa VD, Rojjanasrirat W. Translation, adaptation and validation of instruments or scales for use in cross-cultural health care research: a clear and user-friendly guideline. *J Eval Clin Pract.* 2011;17(2):268–74. x.Epub 2010 Sep 28.
15. Hatcher L, O'Rourke N. Minimal Sample Size Requirement. A Step-by-Step Approach to Using the SAS System for Factor Analysis and Structural Equation Modeling. North Carolina: SAS Institute Inc.; 2014. p. 9.
16. Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, Bouter LM, de Vet HC. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol.* 2007;60(1):34–42. Epub 2006 Aug 24. PMID: 17161752.
17. Streiner DL. Being inconsistent about consistency: when coefficient alpha does and doesn't matter. *J Pers Assess.* 2003;80:217–22.
18. McGraw KO, Wong SP. Forming inferences about some intraclass correlation coefficients. *Psychol Methods.* 1996;1:30–46.
19. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med.* 2016;15(2):155–63. <https://doi.org/10.1016/j.jcm.2016.02.012>. PMID: PMC4913118.
20. Parpala A, Lindblom-Ylänne S, Komulainen E, Litmanen T, Hirsto L. Students' approaches to learning and their experiences of the teaching-learning environment in different disciplines. *Br J Educ Psychol.* 2010;80(Pt 2):269–82. <https://doi.org/10.1348/000709909X476946>.
21. Parpala A, Postareff L. Development and validation of the HowULearn questionnaire for student feedback. *Stud Educational Evaluation.* 2018;56:40–7.
22. Mikkonen J, Ruohoniemi M, Parpala A. Understanding medical students' learning experiences: A cross-cultural analysis. *Med Educ.* 2020;54(11):995–1006.
23. Dyrbye LN, Thomas MR, Shanafelt TD. Systematic review of depression, anxiety, and other indicators of psychological distress among U.S. and Canadian medical students. *Acad Med.* 2006;81(4):354–73.
24. Tempiski P, Bellodi PL, Parpala A, Paro HBMS, et al. Learning approaches and well-being in medical education: A multicenter study. *BMC Med Educ.* 2021;21(1):41.
25. Steinert Y. Developing medical educators: a journey of lifelong learning. *Med Teach.* 2020;42(8):898–904.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

Shaniko Kaleci¹ · Stefania Costi^{1,2,3} · Pierantonio Bellini¹ · Luigi Generali¹ · Antonella Lotti⁵ · Paola Sena¹ · Johanna Chester¹ · Giovanna Amaducci⁴ · Daniela Mecugni^{1,4}

✉ Shaniko Kaleci
Shaniko.kaleci@gmail.com

¹ Department of Surgery, Medicine, Dentistry and Morphological Sciences with Transplant Surgery, Oncology and Regenerative Medicine Relevance (CHIMOMO), University of Modena and Reggio Emilia, Modena, Italy

² University of Modena and Reggio Emilia, Modena, Italy

³ Azienda Unità Sanitaria Locale - IRCCS di Reggio Emilia, Reggio Emilia, Italy

⁴ University of Modena and Reggio Emilia, Reggio Emilia, Italy

⁵ Department of Humanities. Letters, Cultural Heritage, Education Sciences, University of Foggia, Foggia, Italy