



## The Design and Validation of a Self-Assessment Questionnaire of Critical Appraisal of Medical Literature Skills in Medical Students

MONA MLIKA<sup>1,2,3\*</sup>, MD; ABDELMAJID NACEUR<sup>4</sup>, PhD; FAOUZI MEZNI<sup>1,3</sup>, MD; LILIA ZAKHAMA<sup>1</sup>, MD; IHEB LABBENE<sup>1</sup>, MD; MOHAMED JOUINI<sup>1</sup>, MD

<sup>1</sup>University of Tunis El Manar, Faculty of Medicine of Tunis, Tunis, Tunisia; <sup>2</sup>Department of Pathology, Center of Traumatology and Major Burns, Ben Arous, Tunis, Tunisia; <sup>3</sup>Laboratory Research: LR18SP06, Ministry of Health, Tunis, Tunisia; <sup>4</sup>Higher Institute of Education ISEFC, Research Unit ECOTIDI (UR 16E S10), Tunis, Tunisia

### Abstract

**Introduction:** Critical appraisal of medical literature is a challenging step of the evidence-based medicine practice. Many assessment questionnaires have been published in the literature, but they have mainly focused on all the evidence-based medicine practice process. The authors aimed to develop and validate a questionnaire assessing the critical appraisal skills of medical students from the same Faculty.

**Methods:** The questionnaire was developed by item generation through a review of the literature and an expert committee. The questionnaire was validated in terms of content validity and construct validity. Fitness of data for analysis was checked through Kaiser-Meyer-Olkin (KMO) and Bartlett's sphericity. Construct validity was carried out using a principal axis exploratory factor analysis (EFA) with 'varimax' rotation to study the internal structure of the questionnaire and to extract the test major factors. The questionnaire was administrated to a cohort of under and postgraduate medical students (n=84) to evaluate the test reliability and select the best items. The reliability of the questionnaire was assessed by Cronbach's alpha coefficient to evaluate the internal consistency. The correlations between the self-confidence and satisfaction dimension score, the critical thinking dimension score, the learning style dimension score, the Fresno-adapted test scores and the total score were assessed using the Spearman's correlation test.

**Results:** The questionnaire consisted of 31 items. A factorial analysis grouped the items into 3 dimensions that consisted of the self-confidence and satisfaction dimension, the critical thinking dimension and the learning style dimension. Cronbach's alpha accounted for 0.95, CI95% [0.9-1] for the entire questionnaire. The factor analysis explained 79.51% of the variance. The external validity assessment based on a Spearman's correlation study highlighted a weak correlation between the total scores and the critical thinking dimension and the self-perception and satisfaction dimension.

**Conclusion:** In spite of the limitations of this study, mainly the small number of the students recruited, the questionnaire seems to measure with adequate reliability the competences of under and postgraduate medical students.

**Keywords:** Questionnaire, Validation, Medical in literature, Validity, Reliability

\*Corresponding author:  
Mona Mlika, MD;  
Department of Pathology,  
Abderrahman Mami  
Hospital,  
Ariana. 2084, Tunis, Tunisia  
**Tel:** +21-98538862  
**Email:** mouna.mlika@fmt.  
utm.tn

Please cite this paper as:  
Mlika M, Naceur AM,  
Mezni F, Zakhama L,  
Labbene I, Jouini M. The  
Design and Validation  
of a Self-Assessment  
Questionnaire of Critical  
Appraisal of Medical  
Literature Skills in Medical  
Students. J Adv Med  
Educ Prof. 2023;11(2):69-  
79. DOI: 10.30476/  
JAMP.2023.97189.1746.

**Received:** 9 November 2022

**Accepted:** 8 February 2023

## Introduction

Evidence-based medicine (EBM) skills are compulsory in medical practice. Many teaching techniques have been described in the literature that are mainly sustained by cognitivist and behaviourist learning theories. The assessment of the critical appraisal skills varied in the different studies. Some authors assessed these skills using self-assessment or self-confidence questionnaires before and after the training, focusing on cognitive skills. Others, also used questionnaires focusing on the reflexive skills or attitudes using scales like the EPIC scale (1). Moharari RS, et al. used the Consolidated Standards of Reporting Trials (CONSORT) checklist before and after the course in order to assess the students' skills (2). Nieminen P, et al. used a self-administered questionnaire (3). The questionnaire consisted of three different questions assessing the students' knowledge of EBM principles. Shehata GM, et al. assessed critical thinking and attitudes of the learners using a self-administrated questionnaire assessing cognitive skills and reflexive attitudes (4). Stern DT, et al. used the script concordance test concept by using a questionnaire containing structured questions about a specific article and comparing the residents' answers to the tutors' answers (5). Another consensual test is represented by the FRESNO test which assesses EBM skills including critical appraisal competences (6). The authors aimed to design and validate a questionnaire assessing the cognitive, conative, critical thinking, attitudes and satisfaction of the participants to training sessions centred on critical appraisal of medical literature learning. They developed a post-training self-assessment questionnaire that was designed and validated according to the questionnaire developed by Tsang S, et al. (7).

## Methods

The validation of the questionnaire occurred according to the Tsang S, et al's. steps (7). The first phase consisted of the formation of an expert committee and the identification of items for the questionnaire, the second phase involved a pilot test of the first version, the assessment of the internal reliability, an exploratory factor analysis with orthogonal varimax rotation for each item and an external examination of correlations between the attributes and other characteristics of the respondents.

**Participants:** The participants were undergraduate and postgraduate medical students who accepted to be involved in a training workshop centred on the critical appraisal of

medical literature and tutored by the same tutor. The sample size wasn't calculated because of the volunteer character of the participation. The students included fulfilled all the items of the questionnaire and answered to a final test adapted from the Fresno test. The latter is a published test containing 7 short answer questions, 2 mathematical questions and 3 fill-in-the blank questions. It was used to assess the residents' knowledge of EBM principles. We modified the Fresno test in order to focus only on the critical appraisal of the literature. For that reason, 6 short-answer questions were added in relation to a part from an original manuscript dealing with a diagnostic test.

**Item development and selection:** In phase one, a review of the literature was performed by searching for papers and questionnaires related to the assessment of evidence-based practice in general and critical appraisal competencies in particular. Scientific databases (PubMed, Embase) were consulted looking for research published between 2015 and 2022. General descriptors employed were: 'critical appraisal of medical literature' AND 'evidence' AND 'evidence-based medicine' AND 'practice' AND 'knowledge' AND 'skills'. The search was limited to English language.

For the second step, an expert committee was formed in order to develop the items.

**The expert committee:** An expert committee was created in order to conceive the different items. This committee used to meet once a month during 3 months in order to assess the literature and to prepare the different dimensions of the questionnaire. This committee was composed of three full professors who were used to tutor students from different levels and to teach the different principles of the EBM in general with focus on the critical appraisal of medical literature. Nine workshops were organized and during these workshops, 150 students attended the workshops. 84 students completed the questionnaire.

**Identifying dimensionality of construct:** The experts reviewed the literature in order to identify the different dimensions able to influence the practice of critical appraisal of medical literature. Even if, the majority of the studies dealing with the critical appraisal practice were based on satisfaction questionnaires or concerned formative assessment, the experts supposed that learner style, the attitudes, the conative profile, the self-confidence, the critical thinking profile and the satisfaction were more likely to influence the practice of appraising medical literature.

**Description of the initial questionnaire:**

The initial questionnaire conceived by the experts assessed 6 dimensions that consisted of: the learning style of the trainees, the attitudes, the conative dimension, the critical thinking dimension, the self-confidence potential and the satisfaction.

**Learning style:** The learning strategies of medical students focused on their ability to use real cases of patients for their learning, their potential to recall facts related to a new learning by using many exercises or by repeating the information till memorizing it. The aim of this dimension was to differentiate learners who prefer using active pedagogy principles and tend to be autonomous and those who prefer traditional pedagogy and are passive learners. The questions were adapted from the questionnaire used by Nieminen P, et al. in their study about the assessment of critical appraisal competencies among Finnish dental students (3).

**Attitudes dimension:** Assessing the attitude dimension focused on the students' tendency to be interested only in attractive themes or even in necessary ones. The questions related to the attitudes were adapted from the questionnaire of Baum KD, et al. (8).

**Conative dimension:** Assessing the conative domain focused on the students' potential to apply the different principles they learn in clinical settings. For that reason, the students were asked about their motivation to become a doctor: Do they want to become doctor for wealth or for saving lives? The questions used in the conative domain were adapted from the questionnaire published by Maloney LM, et al. (9).

**Critical thinking dimension:** Assessing the critical thinking profile of the students focused on their needs to search for criteria, their sensitivity to context and their tendency to regularly assess their competencies and to correct their learning strategies when needed. The questions were adapted from the questionnaire published by Shehata GM, et al. about critical thinking and attitude of physicians towards EBM in Egypt (4). The questions related to the critical thinking potential of the participants were articulated around its 3 characteristics stipulated by Lipman M. (10). These characteristics consist of searching for criteria, sensitivity to context and self-correction.

**Self-confidence dimension:** The questions related to the self-confidence domain were adapted from the evidence-based practice confidence (EPIC) scale published by Rohwer A, et al. (1). They assessed the students' self-confidence about their capacities to appraise

medical literature after the workshop.

**Satisfaction domain:** The questions assessing the participants' satisfaction were centred on the students' satisfaction concerning the duration of the training or the training environment.

After the identification of the construct's dimensionality, the expert committee developed the items, determined the questionnaire length and reviewed the initial items pool.

**Questionnaire presentation:** For the expert committee, every researcher made a thematic analysis of the data and coded them for quite specific concepts or dimension. A consensus was reached concerning the items and the rating of each item.

**Rating the different dimensions:** As the questions were Likert scaled, the experts agreed that items scored 4 or 3 by the students highlighted high levels related to the dimension assessed, the items scored 0 and 2 pointed out intermediate levels and items scored 1 or 0 were related to low levels. The accurate rating of each dimension is detailed below:

**Learning style:** This dimension contained 4 items. Scores accounting for <8 were correlated to a passive learning strategy. Scores tending to be between [8-12] were correlated to an intermediate learning strategy and scores between [12-16] were correlated to an active learning strategy.

**Attitudes:** This dimension contained 3 items. Scores between [9-12] were considered correlated to a positive attitude towards the critical appraisal practice, scores between [0-6] as correlated to a negative attitude and scores between [6-9] as correlated to an intermediate attitude.

**Self-confidence dimension:** This dimension contained 16 items. The experts considered scores between [48-64] as correlated to a good self-confidence perception, scores between [32-48] as correlated to an intermediate self-confidence perception and scores <32 as correlated to a negative self-confidence perception.

**Conative dimension:** This dimension contained 6 items. The experts considered scores between [18-24] as correlated to a high potential to use the principals of critical appraisal, scores between [12-18] as correlated to an intermediate potential to reproduce the experience and scores between [0-12] as correlated to a negative potential to appraise medical literature.

**Critical thinking dimension:** This dimension contained 4 items. The experts considered scores between [0-8] as correlated to a low critical thinking potential, scores between [8-12] as correlated to an intermediate critical thinking potential and scores between [12-16] as correlated to a high critical thinking potential.

**Satisfaction dimension:** This dimension contained 3 items. The experts considered scores between [9-12] as correlated to a high satisfaction, scores between [6-9] correlated to a mild satisfaction and scores <6 correlated to dissatisfaction.

#### *Validation procedures and data analysis*

**Content validity:** The experts answered clear and easy questions covering all the determinants of critical appraisal practice.

**Pre-Validation study:** Thirty medical students in the third year of medical education (TYME) were asked about their understanding of the reviewed items. After modifying the items according to the students' comprehension and suggestion, the authors administered the first draft to a sample of TYME and continuing medical education (CME) students. The instrument was administered to the different students who attended the same training, tutored by the same tutor and centred on the critical appraisal of medical literature.

**Construct validity:** Kaiser-Meyer-Olkin (KMO) and Barlett's sphericity tests were used to check the fitness of data and the sample adequacy.  $KMO > 0.5$  was related to a good sampling adequacy. Construct validity was carried out using a principal axis exploratory factor analysis (EFA) with 'varimax' rotation. The number of dimensions or items was validated by examining the sorted factor loadings, eigen values and scree plots. Kaiser criterion (eigen values > 1 only once) and the degree of explanation of variability (first main component > 40%) were used to validate unidimensional domains (11). The range of factor loading scale is usually between -1.0 and +1.0 and the experts agreed to accept values above 0.5 (12). Examining the correlations between attributes and the characteristics was used to evaluate the external construct validity. Specifically, we hypothesized that the score of the questionnaire should be highly intercorrelated with the learning style or the critical thinking profile. Besides, we hypothesized that the adapted Fresno test should be correlated to the self-perception-satisfaction profile and learner style, whereas the critical thinking profile should be only modestly intercorrelated.

**Internal consistency:** The internal consistency was evaluated using the Cronbach's alpha coefficient (13). The experts agreed that the accepted value of internal consistence is at least 0.6 to 0.7. If the Cronbach Alpha coefficient value for a question is below 0.6, the experts agreed to delete the item. Because of COVID-19 restrictions, test-retest reliability wasn't administered.

**Revision of the items of the questionnaire:** In this

final stage, the experts could revise items or delete them according to the principal component analysis (PCA) and Cronbach's alpha coefficient (12).

**External validity:** It was assessed using the Pearson's or Spearman's coefficients.

The mean score and standard deviation were used to express the different scores, while frequency was used for categorical variables. The correlations between the self-confidence and satisfaction dimension score, the critical thinking dimension score, the learning style dimension score, the Fresno-adapted test scores and the total score were assessed using the Spearman's correlation test. The Kolmogorov-Smirnov test for normality was applied, and all analyses were performed using IBM Corp. Released 2015. IBM SPSS Statistics for Windows, Version 23.0. Armonk, NY: IBM Corp.

**Research approval:** This study was approved by a research committee of a University Hospital (Ref 01/2022).

**Ethical Consideration:** The present study was conducted according to the principles of the Declaration of Helsinki. Ethical approval for the study was obtained from the ethic committee of a University Hospital (Ref 07/2022). Besides, the participants were made aware of the purpose of the study, the anonymous nature of the purpose, the anonymous nature of the dataset generated and the option to not respond if they so wished. This information served as the basis for an informed consent from each respondent.

## Results

**Preliminary pilot study:** A preliminary pilot study was conducted including 84 students consisting of 68 women and 16 men with a mean age of 24 years (SD: 1.9). 59 students were in the TYME and 24 students were postgraduates.

**Validity of the questionnaire:** This step consisted of assessing the content and construct validity of the questionnaire.

**Content Validity:** To assess the content validity, 2 experts reviewed the final version of the questionnaire and were asked about its adequacy. The experts reviewed the questions and didn't add modifications.

**Construct Validity:** Firstly, the 36 items correlated at least 0.3 with at least one other item, suggesting reasonable factorability. Secondly, the Kaiser-Meyer-Olkin measure of sampling adequacy was 0.87 ( $p < 0.001$ ), suggesting that the factor analysis was appropriate for this data set. Barlett's test of sphericity was significant Chi-square = 2622.285,  $p < 0.001$ ). The diagonals of the anti-image correlation matrix were all over 0.5, supporting the inclusion of each item in the

factor analysis. Finally, the communalities were all above 0.3 further confirming that each item shared some common variance with other items. Visual inspection of the scree plot revealed that the point of inflexion in the plot occurred at the ninth factor, indicating that eight factors should be retained (Figure 1).

Table 1 represents the different factors with the variance of each component. Given these overall indicators, EFA was conducted with all 36 items. The EFA (principal components) using varimax rotation to account for the relationship among the factors yielded an 8-factor structure that explained 79.51% of the variance data.

Figure 2 illustrates the 3 factors retained. They explained 44.16 of the variance of the data. The experts decided to rule out one item (AttQ33) because it couldn't be included in the same axis with the other questions.

Factor 2 contained 7 items dealing with 'critical thinking', explaining 7.73% of the variance of the data. The experts decided to rule out the items ConQ7: 'I want to become a doctor in order to help the society' because it couldn't be part of the critical thinking profile. The learnQ1 question: "for my learning, I need real cases in order to link the theory to the practice" and the item AttQ6 stating that "themes integrating real cases are more interesting"

also centred on the same theme of clinical reasoning and are quite similar so that the experts decided to retain the learnQ1 question because it presented a higher eigen value. ConQ8 and ConQ9 items were ruled out because they showed a high eigen value in more than a factor (respectively, factors 3/factor 4 and factor 7/ factor 8).

Factor 5 contained 2 items explaining 3.72% of the variance of the data. The experts estimated that both questions couldn't be grouped together because LearnQ4 was related to the learner style, and AttQ5 was intimately close to attitudes and behaviourism.

Factors 6 (2 items): 'learner style' contained Learn Q3 and learn Q2 questions.

In conclusion, EFA highlighted 3 factors: Factor 1 (22 items) dealing with 'self-perception and satisfaction profile,' factor 2 dealing with 'critical thinking' and containing 5 items and factor 3 dealing with "learner style". Figure 2 represents the items with their loadings in each factor. Table 2 illustrates the final version of the 31-item questionnaire (Supplementary material). The rating of the different factors was as follows: For factor 1, scores between [66-88] were correlated to a high self-perception and satisfaction potential, scores between [44-66] were correlated to an intermediate potential and scores <44 were correlated to low potential.

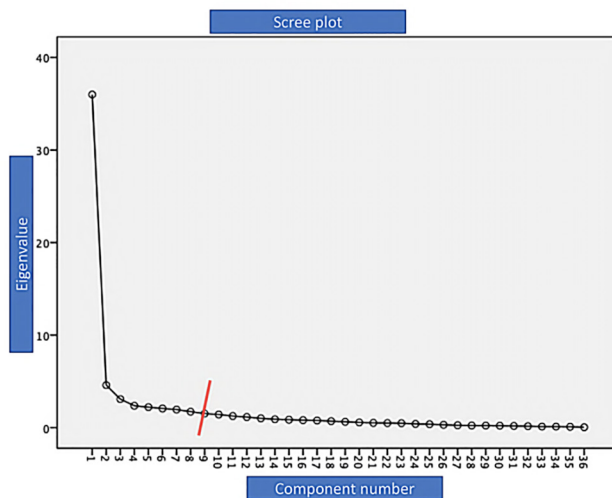


Figure 1: The scree plot graphic showing a point of inflexion occurring at the ninth factor and indicating that eight factors should be retained

Table 1: Variance values of each component

Factors	Total	% of the variance	% Cumulated
1	35.991	51.555	51.555
2	4.591	6.577	58.132
3	3.084	4.417	62.549
4	2.361	3.382	65.931
5	2.214	3.172	69.103
6	2.068	2.962	72.065
7	1.954	2.798	74.863
8	3.248	4.653	79.517

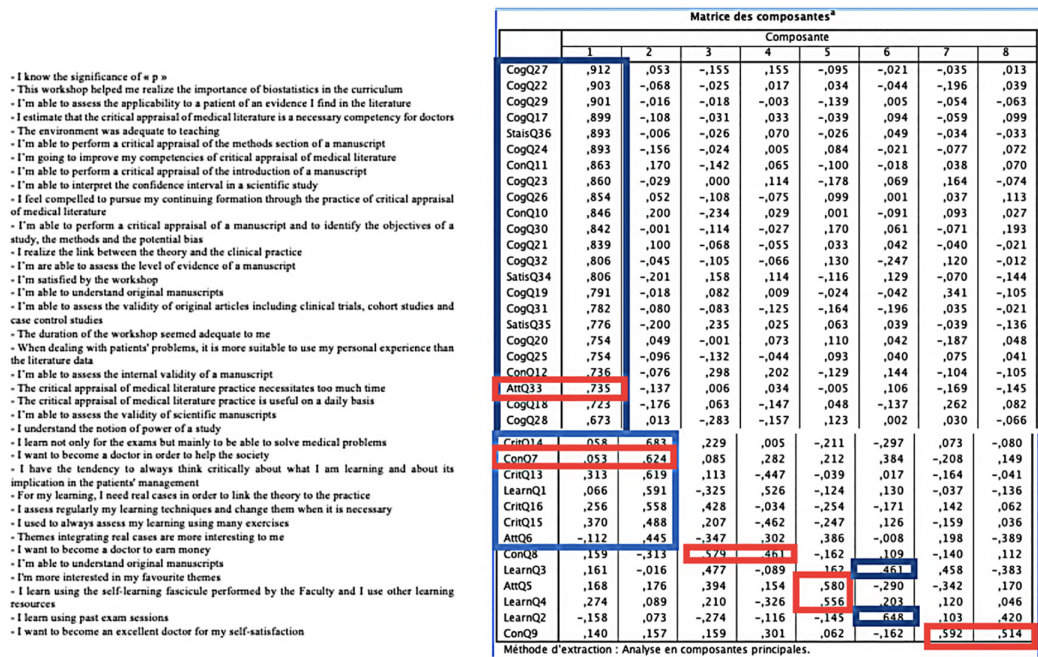


Figure 2: The different loadings per items of the initial questionnaire. The red colour was used for the items ruled out from the initial questionnaire.

**Table 2: The final version of the questionnaire with the different Cronbach's alpha coefficients values**

Items	Rating	Cronbach's alpha coefficients
Self-perception of knowledge and satisfaction		
I know the significance of « p ».	- Totally agree: 4 - Agree: 3 - Not significant: 0 - Disagree: 2 - Totally disagree: 1	0.954
This workshop helped me realize the importance of biostatistics in the curriculum.	- Totally agree: 4 - Agree: 3 - Not significant: 0 - Disagree: 2 - Totally disagree: 1	0.955
I am able to assess the applicability to a patient, of an evidence I find in the literature.	- Totally agree: 4 - Agree: 3 - Not significant: 0 - Disagree: 2 - Totally disagree: 1	0.954
I estimate that the critical appraisal of medical literature is a necessary competency for doctors.	- Totally agree: 4 - Agree: 3 - Not significant: 0 - Disagree: 2 - Totally disagree: 1	0.954
I am able the perform a critical appraisal of the methods section of a manuscript.	- Totally agree: 4 - Agree: 3 - Not significant: 0 - Disagree: 2 - Totally disagree: 1	0.954
I am able the perform a critical appraisal of the introduction of a manuscript.	- Totally agree: 4 - Agree: 3 - Not significant: 0 - Disagree: 2 - Totally disagree: 1	0.956
I am able to interpret the confidence intervals in a scientific study.	- Totally agree: 4 - Agree: 3 - Not significant: 0 - Disagree: 2 - Totally disagree: 1	0.955

I am able to perform a critical appraisal of a manuscript and to identify the objectives of a study, the methods and the potential bias.	- Totally agree: 4 - Agree: 3 - Not significant: 0 - Disagree: 2 - Totally disagree: 1	0.955
I realize the link between the theory and the clinical practice.	- Totally agree: 4 - Agree: 3 - Not significant: 0 - Disagree: 2 - Totally disagree: 1	0.954
I am able to assess the level of evidence of a manuscript.	- Totally agree: 4 - Agree: 3 - Not significant: 0 - Disagree: 2 - Totally disagree: 1	0.950
I am able to understand original manuscripts.	- Totally agree: 4 - Agree: 3 - Not significant: 0 - Disagree: 2 - Totally disagree: 1	0.950
I am able to assess the validity of original articles including clinical trials, cohort studies and case control studies.	- Totally agree: 4 - Agree: 3 - Not significant: 0 - Disagree: 2 - Totally disagree: 1	0.951
When dealing with patients' problems, it is more suitable to use the literature data than my personal experience.	- Totally agree: 4 - Agree: 3 - Not significant: 0 - Disagree: 2 - Totally disagree: 1	0.953
I am able to assess the internal validity of a manuscript.	- Totally agree: 4 - Agree: 3 - Not significant: 0 - Disagree: 2 - Totally disagree: 1	0.955
I am able to assess the validity of scientific manuscripts.	- Totally agree: 4 - Agree: 3 - Not significant: 0 - Disagree: 2 - Totally disagree: 1	0.953
I understand the notion of power of a study.	- Totally agree: 4 - Agree: 3 - Not significant: 0 - Disagree: 2 - Totally disagree: 1	0.955
I feel compelled to pursue my continuing formation through the practice of critical appraisal of medical literature.	- Totally agree: 4 - Agree: 3 - Not significant: 0 - Disagree: 2 - Totally disagree: 1	0.953
I am going to improve my competences of critical appraisal of medical literature	- Totally agree: 4 - Agree: 3 - Not significant: 0 - Disagree: 2 - Totally disagree: 1	0.955
The critical appraisal of medical literature practice doesn't necessitate too much time in routine practice.	- Totally agree: 4 - Agree: 3 - Not significant: 0 - Disagree: 2 - Totally disagree: 1	0.953
I am satisfied by the workshop.	- Totally agree: 4 - Agree: 3 - Not significant: 0 - Disagree: 2 - Totally disagree: 1	0.954

The duration of the workshop seemed adequate to me.	- Totally agree: 4 - Agree: 3 - Not significant: 0 - Disagree: 2 - Totally disagree: 1	0.950
The environment was adequate to teaching.	- Totally agree: 4 - Agree: 3 - Not significant: 0 - Disagree: 2 - Totally disagree: 1	0.951
<b>Critical thinking</b>		
I have the tendency to always think critically about what I am learning and about its implication in the patients' management.	- Totally agree: 4 - Agree: 3 - Not significant: 0 - Disagree: 2 - Totally disagree: 1	0.950
I learn not only for the exams but mainly to be able to solve medical problems.	- Totally agree: 4 - Agree: 3 - Not significant: 0 - Disagree: 2 - Totally disagree: 1	0.951
I used to always assess my learning using many exercises.	- Totally agree: 4 - Agree: 3 - Not significant: 0 - Disagree: 2 - Totally disagree: 1	0.950
I assess regularly my learning techniques and change them when it is necessary 16.	- Totally agree: 4 - Agree: 3 - Not significant: 0 - Disagree: 2 - Totally disagree: 1	0.951
I want to become a doctor in order to help the society.	- Totally agree: 4 - Agree: 3 - Not significant: 0 - Disagree: 2 - Totally disagree: 1	0.949
For my learning, I need real cases in order to link the theory to the practice.	- Totally agree: 4 - Agree: 3 - Not significant: 0 - Disagree: 2 - Totally disagree: 1	0.952
<b>Learning style</b>		
I learn using past exam sessions.	- Totally agree: 4 - Agree: 3 - Not significant: 0 - Disagree: 2 - Totally disagree: 1	0.951
I learn my lessons by repeating them many times.	- Totally agree: 4 - Agree: 3 - Not significant: 0 - Disagree: 2 - Totally disagree: 1	0.951

Twenty-two students obtained low level scores, 34 students obtained intermediate level scores and 28 students obtained high level scores.

For factor 2, scores between [15-20] were correlated to a high critical thinking potential, scores between [10-20] were correlated to an intermediate critical thinking potential and scores <10 were correlated to a low critical thinking potential. Ten students had low critical thinking potential, 31 had intermediate critical thinking potential and 42 had a high critical thinking potential.

For factor 3, scores between [6-8] were correlated to an active learning strategy, scores between [4-6] were correlated to an intermediate learning strategy and scores <4 were correlated to a passive learning strategy. Forty-six students were considered passive learners, 37 students were considered intermediate learners and 1 student was considered as an active learner. The different details of the students concerning the scoring of the questionnaire, the different factors and the Fresno-adapted test are illustrated in Table 3.



**Table 3:** The different scores of the students

	Self-confidence and satisfaction	Critical thinking	Learner style	Total score	Fresno-adapted test
Mean±SD	49.25±2.7	14.02±4.35	3.27±1.23	83.65±29.59	3.96±2.29
Extremes	(0-83)	(0-20)	(0-6)	126-83.65	0-7.5
High level	28 (32.6%)	42 (48.8%)	1 (1.2%)		
Intermediate level	34 (39.5%)	31 (36%)	37 (43%)		
Low level	22 (25.6%)	10 (11.6%)	46 (53.5%)		

**Table 4:** The correlations between the self-confidence and satisfaction dimension score, the critical thinking dimension score, the learning style dimension score, the Fresno-adapted test scores and the total score

		Fresno-adapted test	Self-confidence/satisfaction	Critical thinking	Learning style	Total score
Fresno-adapted test	Spearman coefficient		0.113	0.14	-0.32	0.144
	P-value		0.300	0.180	0.770	0.190
Self-confidence/satisfaction	Spearman coefficient	0.113		0.364	0.034	0.965
	P-value	0.305		0.001	0.757	0.001
Critical thinking	Spearman coefficient	0.145	0.364		-0.016	0.556
	P-value	0.180	0.001		0.880	0.001
Learning style	Spearman coefficient	-0.32	0.034	-0.016		0.079
	P-value	0.770	0.750	0.880		0.474
Total score	Spearman coefficient	0.144	0.965	0.556	0.079	
	P-value	0.191	0.001	0.001	0.474	

**Reliability of the questionnaire:** Assessing the reliability of a questionnaire depends on determining its internal consistency. The Cronbach's alpha of the different domains indicated an adequate internal consistency, 0.956. The Cronbach's alpha values per item are displayed in Table 2.

Correlation between the questionnaires scores and the Fresno test scores was performed in order to assess the construct validity. Non parametric tests were used because of the absence of normality of the different data. There was no correlation between the Fresno test and the total score or the different items of the questionnaires ( $p>0.05$ ). The total score was correlated to the self-confidence and satisfaction and the critical thinking dimensions. It wasn't correlated to the learning style. All the details are displayed in Table 4.

## Discussion

Our results highlighted the different steps of validation of a questionnaire related to the critical appraisal of medical literature skills of medical students. The steps used in the methods section were those published by Tsang S, et al. and consisted of establishing an expert committee, identifying dimensionality of the construct, determining questionnaire format, the items

format and the questionnaire length, reviewing the items, then performing a preliminary pilot test study to assess the reliability and the validity of the instrument (7). Other methods of validation have been reported in the literature such as the Delphi method, which defined accurate steps of content validity consisting of 3 rounds to analyse the results by the experts followed by the steps of validation and assessment of the reliability and the validity of the instrument (14). Even if, the different methods reported in the literature may vary in some details, the main steps consist of a validation of a content by a focus group followed by a pilot test study including a sufficient sample size to assess the reliability and the validity of the instrument. Concerning the sample size, unlike clinical trials, no rules exist to determine it. The sample size has to be as large as possible with varying respondents to questions ratios reported in the literature from 5:1 to 10:1, 15:1 or 30:1 (12, 15, 16). In our study, 84 respondents were included. The initial questionnaire contained 36 items centred on 6 themes: learner style, self-perception of knowledge, conative profile, critical thinking profile, attitudes and satisfaction. The EFA highlighted 3 major factors consisting of the self-perception and satisfaction dimension, the learning style dimension and the critical thinking dimension. In the literature, many

dimensions have been assessed including self-perception of knowledge, satisfaction, attitude towards EBM practice, attitude to research scale, reflexive attitude, limits to find and appraise evidence, change in practice, evidence-based culture or evidence-accessing methods (14, 17-20). Besides, the questionnaires varied from Likert-scale questions questionnaires to questionnaires integrating multiple-choice questions or dichotomal answers (17, 21). Assessing critical thinking profile according to its major characteristics which are the sensitivity to context, searching for criteria or self-correction haven't been reported in the literature. According to our questionnaire's results, the majority of the students (42/84) had a high critical thinking profile. On the other hand, the majority of the students (46/84) were passive learners. The majority of the students included were in the TYME which represents the onset of the clerkship period. The fact that the students may have mainly background questions during this period may explain their passive learning style and lack of autonomy and self-directed learning. Besides, our results didn't show a correlation between the Fresno-adapted test score and the scores attributed to the different factors and there was no correlation between the Fresno-adapted test and the total scores. This finding may be explained by the quality of the training technique used during the workshop.

The strength of this study consisted of the validation of a questionnaire centred on the critical appraisal practice among medical students. The validity and reliability of the questionnaire were highlighted by the KMO test, the test of sphericity and the coefficient alpha Cronbach. This study has some limitations. First, the participants included in this study belonged to the same university. Therefore, generalizing the testing to our universities is suitable. Second, the convergent and discriminant validity of the research tool weren't assessed. Finally, the test-retest reliability method wasn't evaluated and it is necessary to examine the consistency over time.

### Conclusion

Our study was the first study about the validation of a questionnaire centred on the practice of critical appraisal of the medical literature in our country. In spite of its limitations, its further use in other levels other than the TYME or CME will improve its validity and reliability.

### Acknowledgement

We do thank you all the students of the Faculty of Medicine who participated voluntarily

in this study.

### Authors' Contribution

M.M designed the study and wrote the search strategy and A.N performed the literature search. Data acquisition and Data analysis were done by M.M and then revived by A.N, F.M, L.Z, I.L, M.J. M.M prepared the manuscript. All authors contributed to the discussion, read and approved the manuscript and agree to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated resolved.

**Conflict of Interest:** None declared.

### References

1. Rohwer A, Young T, van Schalkwyk S. Effective or just practical? An evaluation of an online postgraduate module on evidence-based medicine (EBM). *BMC Med Educ.* 2013;13(1):77.
2. Moharari RS, Rahimi E, Najafi A, Khashayar P, Khajavi MR, Meysamie AP. Teaching critical appraisal and statistics in anesthesia journal club. *QJM.* 2009;102(2):139-41.
3. Nieminen P, Uma E, Pal S, Laitala ML, Lappalainen OP, Varghese E. Information Retrieval and Awareness about Evidence-Based Dentistry among Dental Undergraduate Students-A Comparative Study between Students from Malaysia and Finland. *Dent J (Basel).* 2020;8(3):103.
4. Shehata GM, Zaki A, Dowidar NL, El-Sayeda I. Critical thinking and attitude of physicians toward evidence-based medicine in Alexandria, Egypt. *Journal of the Egyptian Public Health Association.* 2015;90(3):115-20.
5. Williams BC, Stern DT. Exploratory Study of Residents' Conceptual Framework for Critical Appraisal of the Literature. *Teach Learn Med.* 1997;9(4):270-5.
6. Rice J. Fresno Test of Evidence Based Medicine Grading Rubrics (Form A). USA: Docplayer; 2017.
7. Tsang S, Royse CF, Terkawi AS. Guidelines for developing, translating, and validating a questionnaire in perioperative and pain medicine. *Saudi Journal of Anaesthesia.* 2017;11:S80-9.
8. Baum KD. The Impact of an Evidence-Based Medicine Workshop on Residents' Attitudes towards and Self-Reported Ability in Evidence-Based Practice. *Med Educ Online.* 2003;8(1):4329.
9. Maloney LM, Marshall RT, Werfel PA, Johnson SE. Using a Journal Club Series to Introduce Paramedic Students to Research Fundamentals and Critical Appraisal of Medical Literature. *Prehosp Disaster Med.* 2019;34(4):449-53.
10. Lipman M. Critical thinking: What can it be?. *Analytic Teaching.* 1995;8:1-8
11. Kaiser HF. The varimax criterion for analytic rotation in factor analysis. *Psychometrika.* 1958;23(3):187-200.
12. Aithal A, Aithal PS. Development and Validation of Survey Questionnaire & Experimental Data-A Systematical Review-based Statistical Approach

- [Internet]. 2020 [Cited 13 Jun 2020]. Available from: <https://ssrn.com/abstract=3724105>.
13. Feldt L. A test of the hypothesis that Cronbach's alpha or Kuder-Richardson coefficient twenty is the same for two tests. *Psychometrika*. 1969;34(3):363–73.
  14. Ruano ASM, Motter FR, Lopes LC. Design and validity of an instrument to assess healthcare professionals' perceptions, behaviour, self-efficacy and attitudes towards evidence-based health practice: I-SABE. *BMJ Open*. 2022;12(4):e052767.
  15. MacCallum RC, Widaman KF, Zhang S, Hong S. Sample size in factor analysis. *Psychological methods*. 1999;4(1):84.
  16. Malterud K, Siersma VD, Guassora AD. Sample Size in Qualitative Interview Studies: Guided by Information Power. *Qual Health Res*. 2016;26(13):1753–60.
  17. Ruzafa-Martinez M, Lopez-Iborra L, Moreno-Casbas T, Madrigal-Torres M. Development and validation of the competence in evidence based practice questionnaire (EBP-COQ) among nursing students. *BMC Med Educ*. 2013;13(1):19.
  18. Hendricson WD, Rugh JD, Hatch JP, Stark DL, Deahl T, Wallmann ER. Validation of an Instrument to Assess Evidence-Based Practice Knowledge, Attitudes, Access, and Confidence in the Dental Environment. *J Dental Educ*. 2011;75:131-9.
  19. Patelarou A, Schetaki S, Giakoumidakis K, Lialiou P, Patelarou E. Validation of the Evidence-Based Practice Competence Questionnaire for Nursing Students: A Cross-Sectional Study in Greece. *Nurs Rep*. 2021;11(4):765–74.
  20. Panczyk M, Belowska J, Zarzeka A, Samoliński Ł, Zmuda-Trzebiatowska H, Gotlib J. Validation study of the Polish version of the Evidence-Based Practice Profile Questionnaire. *BMC Med Educ*. 2017;17(1):38.
  21. Capraş RD, Drugan TC, Bolboacă SD. Development and validation of a questionnaire to assess evidence based-practice. In: 2017 E-Health and Bioengineering Conference, EHB 2017. USA: Institute of Electrical and Electronics Engineers Inc.; 2017. p. 129–32.