Development, Validation and Reliability of an Open Book Examination Perception Questionnaire-I: A Cross-sectional Study

VEENA A BEMBALGI[1], PRASHANT B MUKKANNAVAR[2], CASSINDRA C CARDOZ[3], POOJA M SUBLOK[4], SHIVANI S PATIL[5]

ABSTRACT

Introduction: Open Book Examinations (OBEs) are being employed in education, presenting unique opportunities, and challenges to both learners and educators. However, the perception and implication of learning experience remains unanswered.

Aim: To assess the content and face validity and reliability of a pretest Open Book Examination Perception Questionnaire-I (OBEQ-I).

Materials and Methods: The present cross-sectional study was conducted in four phases at KLE College of Physiotherapy, Hubballi, Karnataka, India, from July 2024 to July 2025. The first phase included extensive review of literature and forming the first set of questions (n=24). In the second phase, the questionnaire was shared through a Google form with nine validators and their responses on relevance, content, domain specificity, and suggestions were collected. Based on the recommendations, some questions were deleted/reframed (n=6). The third phase comprised of sharing the revised questionnaire with 18 questions through a Google form with 36 physiotherapy students for rating the items for comprehension and clarity. In the fourth phase, Content Validity Index (CVI) for Item level Content Validity (I-CVI) and Scale Level Content Validity (S-CVI/Ave), Face Validity Index Per Item Wise (I-FVI) and scale wise (S-FVI), and reliability were calculated.

Results: I-CVI for relevance of all items scored 1, except item numbers 7, 11, 14, and 20, which scored 0.88. S-CVI/Ave for relevance was 0.88. I-CVI for content was 1 for all items except item 14, which scored 0.88. S-CVI/Ave for content was 0.99. I-FVI scores for comprehension ranged from 1 to 0.8 and S-FVI/Ave for comprehension was 0.97. I-FVI for clarity ranged between 1 to 0.9 and the S-FVI/Ave score was 0.98.

Conclusion: The 18-item pretest OBEQ-1 revealed an appropriate level of content and face validity based on I-CVI, S-CVI/Ave, I-FVI, S-FVI/Ave scores and suggested an impactful consideration following the analysis of reliability of scale.

INTRODUCTION

The OBEs are increasingly used in education, presenting unique opportunities, and challenges to both learners and educators [1]. The University Grants Commission (UGC) emphasises using a variety of assessment methods, including open book exams, to evaluate students [2]. The National Education Policy (NEP) 2020 suggests using different methods to assess student's performance [3]. As educators aim to improve their assessment methods and improve the learning experience, there is an increasing need to explore the dynamics of OBEs and assess its implementation [4].

The OBEs allow the use of textbooks, notes, journals, and/or other reference materials [5]. The reported benefits of OBEs include reduced anxiety, emphasis on understanding concepts and principles instead of memorisation, encouragement of "a deeper engagement with the course material on the part of the students," and use of the material at a higher Bloom's level [5-7].

The OBEs better simulate real life experiences for health practitioners, who often rely on resources when making clinical decisions [8,9]. This approach of practising and making decisions is not new for physiotherapists. However, OBEs have not been explored in physiotherapy education, and students were assumed to be unfamiliar toward this examination pattern. Therefore, it is important to examine the effectiveness of OBEs in physiotherapy education and to identify any misconceptions and misunderstandings that might hinder behaviour change before implementing OBEs.

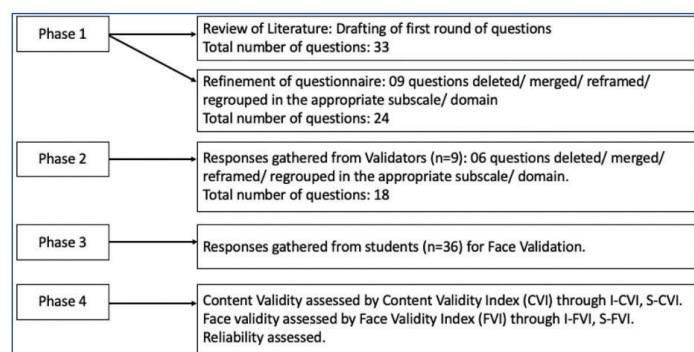Knowledge, Attitude, and Practice (KAP) surveys provide a structured approach for gathering comprehensive insights into students' perceptions of OBEs and are designed to be a representative of a target population [10]. This aims to uncover what is known (knowledge), believed (attitude), and done (practised) in the context of the topic of interest. Information is gathered using semi-structured or (more commonly) structured questionnaires that are self- administered or administered by interviewers. Both qualitative and quantitative data may be collected [11]. KAP surveys helps educators identify knowledge gaps, understand attitudes, evaluate practices, inform pedagogical strategies, promote continuous improvement, and encourage dialogue. This can enable educators to optimise the implementation of OBEs and improve student learning experiences [12].

Zagury-Orly I and Durning SJ, in their review on assessing OBEs in medical education emphasised on the importance of exploring how learners perceive OBEs using well-designed surveys and qualitative research methods [5]. Jervis CG and Brown LR highlighted the need for developing a pretest and post-test treatment of attitude [13]. To the best of our knowledge, no standard questionnaire has been developed yet to capture students' perception of learning in terms of clarity and comprehension through OBEs. Moreover, it is necessary to understand the students' views. This requires the development and standardisation of a feedback questionnaire that can provide valuable insights into the perceptions, experiences, and outcomes associated with OBEs. Therefore, the present study aimed to develop, validate, and assess reliability of OBEQ-I.

MATERIALS AND METHODS

The present cross-sectional study was conducted at KLE College of Physiotherapy, Hubballi, Karnataka, India from July 2024 to July 2025. The ethical clearance for the development of questionnaire

and its testing was obtained from JGMM Medical College Institution Ethics Committee (JGMMMCIEC) with reference number JGMMMCIEC/43/2024. This questionnaire was developed in four phases [Table/Fig-1].

Phase 1	Review of Literature: Drafting of first round of questions Total number of questions: 33
	Refinement of questionnaire: 09 questions deleted/ merged/ reframed/ regrouped in the appropriate subscale/ domain Total number of questions: 24
Phase 2	Responses gathered from Validators (n=9): 06 questions deleted/ merged/ reframed/ regrouped in the appropriate subscale/ domain. Total number of questions: 18
Phase 3	Responses gathered from students (n=36) for Face Validation.
Phase 4	Content Validity assessed by Content Validity Index (CVI) through I-CVI, S-CVI. Face validity assessed by Face Validity Index (FVI) through I-FVI, S-FVI. Reliability assessed.

[Table/Fig-1]: Study procedure flowchart.

Study Procedure

Phase 1: Development of pretest OBEQ-I: The first phase involved extensive review of literature. Thorough literature review was done using PubMed, Google Scholar, and grey literature. Based on information gathered from previous studies [4-6], it was decided by the authors to form a pretest questionnaire to assess the perception of the students about OBEs before they appear for one. Accordingly, a pretest questionnaire was first developed which included general questions and questions in three domains namely, cognitive, affective and psychomotor domain.

A set of 33 questions were framed and after deliberations with experts in the research panel of the institution, nine questions were deleted. A set of 24 questions was agreed to for further assessment. As per the panel recommendation, the domain names were changed to knowledge, attitude and practice for the ease of understanding.

Phase 2: Content validity of OBEQ-I: It consisted of selecting validators for content validation of the questionnaire. A total of 9 validators were agreed on [14]. A judgmental sampling method was used to choose the validators for the content validation of the questionnaire. Validators chosen were experts in the field of education, English Professors, Physiotherapists, and Psychologists based on their expertise and experience in their fields of education.

A Google form was prepared for the validators in which their consent was taken and they were asked to mark the relevance, content and domain specificity and any other suggestions that they deemed valid for each item in the questionnaire. A 5-point Likert Scale was provided for each question to mark the relevance of the question and like-wise to assess the content of the question. The subscale/domain specificity was a Yes/No response to assess if the validator felt the question belonged to the right subscale/domain. Based on the responses and suggestions received, six questions were reviewed and were either deleted/merged/regrouped in the appropriate subscale/domain in the questionnaire. The redrafted questionnaire had a total of 18 questions with questions in general, knowledge, and practice subscale/domain.

Phase 3: Face validation of the pretest OBEQ-I: The final draft of 18 questions was given to 36 students from 2nd year Physiotherapy program in Google forms. A convenience sampling method was used. They were asked to mark the comprehensibility in terms of ease of understanding the question and clarity of the question in terms of language or interpretation on a 5-point Likert scale [15,16]. The acceptable CVI was fixed at 0.8 [17]. All ratings with response on Likert scale of 3 and above were coded as 1 and below 3 were coded as 0 for the sake of analysis. The reliability was analysed using Cronbach's alpha reliability analysis with a threshold of $\alpha \geq 0.7$ defined as acceptable reliability [15].

Phase 4: Calculation of content and face validity; and reliability: Based on the responses received from the validators and students (respondents), content validity and face validity were calculated respectively.

STATISTICAL ANALYSIS

Microsoft Excel was used to tabulate the data and conduct all statistical tests. CVI was calculated for I-CVI and S-CVI/Ave for relevance and content of each item in the questionnaire. The validators were requested to score the relevance of the item on a 5-point Likert scale with 1-not at all relevant to 5-extremely relevant. Similarly, they were requested to score for content of the question on a 5-point Likert scale with 1-very poor content to 5- excellent content. S- CVI/Ave, was the average of I-CVI scores for all items/questions in the questionnaire. Recoding of the data was done as 1 for all scores by validators between 3-5 and 0 for all scores by validators between 1-2 (1=very poor, 2=poor) [15]. All calculations were done manually on an excel sheet based on the formulas recommended by Yusoff MSB 2019 [14]. Face validity was calculated for clarity and comprehension of the questions from responses collected from 36 undergraduate 2nd year Physiotherapy students of KLE College of Physiotherapy, Hubli, Karnataka, India. The respondents rated each question for comprehensibility on a 5-point likert scale (1=not understood, 2=slightly understood, 3=Moderately understood, 4=understood, 5=completely understood) and clarity on a 5-point Likert scale (5=content is not clear at all, 4=content is not clear, 3=content is moderately clear, 4=content is clear, 5=content is very clear). Face Validity Index was calculated per item-wise (I-FVI) and scale-wise (S-FVI). Calculation of I-FVI and S-FI was done manually based on recommendations by Yusoff MSB, 2019 [14]. The reliability was analysed using Cronbach's alpha, exploratory factor analysis, confirmatory factor analysis, and convergent and discriminatory validity analysis.

RESULTS

All the validators belonged to academic background of Physiotherapy, English Professors, and Psychologists. The mean age of the validators with a teaching experience of the experts from the field of education is mentioned in [Table/Fig-2].

Sociodemographic variables	Mean±SD
Age	50.33±4.4 years
Teaching experience	13.6±2.1 years

[Table/Fig-2]: Sociodemographic variables of experts.

Content Validity in terms of I-CVI, S-CVI was calculated for relevance and content. I-CVI of all items for relevance were 1 except for 7th item, 11th, 14th and 20th item, which received scores of 0.88 but were in acceptable CVI ranges. S-CVI/Ave for relevance was 0.88 and within acceptable scores. I-CVI for content was one for all items except 14th item, which was 0.88. S-CVI/Ave for content of all the items was 0.99 which was well within the acceptable score.

Face validity in terms of I-FVI, S-FVI were calculated for comprehension and clarity of each question. The acceptable FVI was fixed at 0.8 [16]. All the responses in the Likert scale of three and above were coded as one and below three were coded as 0 for the sake of analysis and S-FVI/Ave for all items for comprehension and clarity. I-FVI scores for comprehension ranged between 1 to 0.8 for all items and S-FVI/Ave for comprehension was 0.97 and was within acceptable score. I-FVI for clarity ranged between 1 to 0.9 for all items and the S- FVI/Ave score was 0.98. The reliability was analysed using Cronbach's alpha reliability as shown in [Table/Fig-3].

The exploratory factor analysis was done using Kaiser-Meyer-Olkin (KMO) test and Bartlett's test. A KMO of 0.82 reported a meritorious factor with a very significant (p<0.0001) through Bartlett's test. There were three factors that were used for determining the suitability of content. The factor 1 referred to the answer quality with α=0.72-

0.85 which was found in Q8, Q9, Q10; factor 2 referred to the stress/performance with α=0.65-0.78 which was found in Q13, Q14, and Q15; and factor 3 referred resource use and stated with α=0.68-0.81which was specified in Q6a, Q6b, Q7.

Cronbach's alpha reliability analysis			
Threshold: α≥0.7=acceptable reliability			
Subscale/domain	Items	α	Interpretation
Familiarity with exams	Q1a (OBE), Q1b (Closed), Q1c (Oral)	0.68	Slightly below threshold
OBE understanding	Q2a (Assignments), Q2b (Ref. Materials), Q3a (No experience), Q3b (Type), Q4 (Preparation), Q5 (Materials)	0.72	Acceptable
Resource utilisation	Q6a (Textbooks), Q6b (Digital), Q7 (Locate Info)	0.81	Excellent
Answer quality	Q8 (Detailed answers), Q9 (Summarising), Q10 (Synthesise)	0.84	Excellent
Time management	Q11 (Time pressure), Q12 (Search time)	0.73	Acceptable
Stress and performance	Q13 (Stress), Q14 (Performance), Q15 (Preference)	0.69	Borderline
Academic integrity	Q16 (Cheating), Q17 (Anxiety)	0.75	Acceptable

[Table/Fig-3]: Reliability analysis of the content.

The confirmatory factor analysis is demonstrated in [Table/Fig-4] which interpreted good extent of observed frequencies with acceptable Comparative Fit Index (CFI), good Root Mean-Square Error of Approximation (RMSEA), and excellent Standardised Root Mean Square Residual (SRMSR) supporting a good fit of scale. The convergent and discriminatory validity is demonstrated in [Table/Fig-5] which stated a strong convergent validity for answer quality and acceptable academic integrity but borderline for time management.

Index	Value	Threshold	Interpretation
χ^2/df	2.1	<3.0	Good
CFI	0.93	≥0.90	Acceptable
RMSEA	0.06	≤0.08	Good
SRMSR	0.04	≤0.08	Excellent

[Table/Fig-4]: Confirmatory factor analysis of the content.
Where, df=degree of freedom; CFI: Comparative fit index; RMSEA: Root mean square error of approximation; SRMR: Standardised root mean square residual

Subscale/domain	AVE	CR	Interpretation
Answer quality	0.62	0.84	Strong convergent validity
Time management	0.51	0.73	Borderline (AVE <0.5 is risky)
Academic integrity	0.58	0.75	Acceptable

[Table/Fig-5]: Convergent and discriminant validity.
Where, AVE: Average variance extracted, CR: Composite reliability

DISCUSSION

The OBEQ-I was created as an initial step for examining and implementing OBEs in the curriculum for Physiotherapy. The study included development, validation, and reliability analysis of the questionnaire. The validators were from diverse field to capture the nuances of the questionnaire from their diverse expertise and field of practice. The validators selected had broad professional background, which ensured that the questions included in the questionnaire were meaningful, relevant, and critically examined. An engineering professor was approached on purpose since OBEs are conducted on a regular basis in the said field of education. Therefore, an input from that field of education would be valuable. The Psychologists were included to understand the mindset of students before they appear an OBE. Their input on development of the questionnaire was significant.

Recommended number of validators for content validation is a minimum of six panels of validators [14]. A panel of nine validators

was included to improve the robustness of the questionnaire. The content was validated for relevance using both I-CVI and S-CVI. Following the guidelines given by Yusoff MSB (2019) [14], 0.8 was considered as the acceptable threshold for CVI for nine validators. The I-CVI for relevance indicated that most items achieved a perfect score of 1, except for items 7, 11, 14, and 20, which had scores of 0.88. Despite these deviations, all items were within the acceptable CVI range. The overall S-CVI/Ave for relevance was 0.88, which was within the threshold value.

Similarly, the I-CVI for content of all items scored 1, except for item 14, which scored 0.88. The reduced score suggests some concerns about its relevance; however, it was still within acceptable limits and hence was included in the questionnaire after consensus was gathered from the research panel. These results indicate the robustness of the instrument in terms of relevance and content validity. Although some items have slight deviations in scores, these do not significantly reduce the overall validity of the scale but highlight areas for improvement. A study by Anthoine E et al., aimed at determining the sample size for newly developed outcome measures found that there is no gold standard for the recommended sample size to validate a newly developed instrument [18]. According to Yusoff MSB, 2019, the minimum recommended number of respondents for response or face validation is ten [14]. Although, a convenient sample was used, the students were not exposed to any formal OBE. Thus, assuring a fair response rating from the respondents.

The I-FVI and S-FVI/Ave values for all items were within the acceptable range, indicating that all items were clear and understandable. Thus, the respondents understood the questionnaire well, which was effective in capturing the intended constructs. The questionnaire is such designed that it can used to capture students KAP not necessarily from physiotherapy field alone but from all other fields of education. Though, Physiotherapy students participated in the study for face validation, face validation for the same questionnaire can be done from other fields of education ranging from preuniversity courses to diploma and graduate courses in all disciplines. It is also possible that the feedback received from students through this questionnaire can help educators reform their evaluation processes and include OBEs as a formal assessment method in the curriculum. The high CVI and FVI scores suggest that the instrument is both valid and reliable for assessing the targeted constructs. The few items that scored slightly below one, particularly item 14, indicate areas where the instrument can be refined. Future studies might consider revising these items to enhance their clarity and relevance further.

Limitation(s)

The study collected data from the students which might result in the subjective understanding of the comprehension and might lead to bias. The sample size can be increased for considering the larger impact.

CONCLUSION(S)

Overall, the findings support the use of this instrument in educational settings, particularly for evaluating the perception of open book exams. The rigorous validation process undertaken in the present study provides confidence in the instrument's ability to accurately measure the intended constructs, making it a valuable tool for educators and researchers alike. By ensuring high levels of content and face validity, a better understanding and improvement in the assessment methods used in educational contexts could be obtained, ultimately enhancing the learning experiences and outcomes for students. Also, a wide scope of population can be recruited in future to check for the perception of the scale from different educational sectors. A predictive validity and responsiveness to change can be investigated to establish whether OBEQ-I scores predict subsequent behaviours or outcomes. Translation, cultural adaptation, and linguistic validation can be done in future.

REFERENCES

[1] Eldakhakhny B, Alamoudi AA, Gad H, Almoghrabi Y, Shamrani T, Daghistani H, et al. Introducing group open-book exams as a learning and assessment strategy in the clinical biochemistry course for medical students. Cureus. 2024;16(1):e51792.

[2] Marg BSZ. University Grants Commission.

[3] About National Education Policy | Government of India, Ministry of Education [Internet]. [cited 2025 Jul 7]. Available from: https://www.education.gov.in/en/nep/about-nep.

[4] Durning SJ, Dong T, Ratcliffe T, Schuwirth L, Artino AR, Boulet JR, et al. Comparing open-book and closed-book examinations: A systematic review. Acad Med J Assoc Am Med Coll. 2016;91(4):583-99.

[5] Zagury-Orly I, Durning SJ. Assessing open-book examination in medical education: The time is now. Med Teach. 2021;43(8):972-73.

[6] Cox LA, Hariharan RT, Aung YYM. Open book examinations: A student's perspective. Med Teach. 2021;43(6):724.

[7] Gharib A, Phillips W, Mathew N. Cheat Sheet or open-book? A comparison of the effects of exam types on performance, retention, and anxiety [Internet]. Vol. 2, Online Submission. 2012 Aug [cited 2025 Jul 7] p. 469–78. Available from: https://eric.ed.gov/?id=ED537423.

[8] Eilertsen TV, Valdermo O. Open-book assessment: A contribution to improved learning? Stud Educ Eval. 2000;26(2):91-103.

[9] Kruidering-Hall M, Tuan RL. Information overdose: Student performance and perceptions of pharmacology resources on exams. Pharmacol Res Perspect. 2023;11(3):e01087.

[10] Zarei F, Dehghani A, Ratanasiri A, Ghaffari M, Raina SK, Halimi A, et al. ChecKAP: A checklist for reporting a Knowledge, Attitude, and Practice (KAP) study. Asian Pac J Cancer Prev APJCP. 2024;25(7):2573-77.

[11] Andrade C, Menon V, Ameen S, Kumar Praharaj S. Designing and conducting knowledge, attitude, and practice surveys in psychiatry: Practical guidance. Indian J Psychol Med. 2020;42(5):478-81.

[12] Heck OC, Ormiston H, Husmann P. Utilizing KAP in schools: An evaluation of educators' and staff knowledge, attitudes, and practices related to trauma. J Child Adolesc Trauma. 2023;16(4):889-901.

[13] Jervis CG, Brown LR. The prospects of sitting "end of year" open book exams in the light of COVID-19: A medical student's perspective. Med Teach. 2020;42(7):830-31.

[14] Yusoff MSB. Department of Medical Education, School of Medical Sciences, Universiti Sains Malaysia, MALAYSIA. ABC of Content Validation and Content Validity Index Calculation. Educ Med J. 2019;11(2):49-54.

[15] Park DI. Development and Validation of a Knowledge, Attitudes and Practices Questionnaire on COVID-19 (KAP COVID-19). Int J Environ Res Public Health. 2021;18(14):7493.

[16] Rahman MA, Yusoff MSB, Roslan NS, Mohammad JAM, Ahmad A. Development and validation of the medical professionals resilience scale. BMC Health Serv Res. 2021;21(1):482.

[17] Halek M, Holle D, Bartholomeyczik S. Development and evaluation of the content validity, practicability and feasibility of the Innovative dementia-oriented Assessment system for challenging behaviour in residents with dementia. BMC Health Serv Res. 2017;17(1):554.

[18] Anthoine E, Moret L, Regnault A, Sébille V, Hardouin JB. Sample size used to validate a scale: A review of publications on newly-developed patient reported outcomes measures. Health Qual Life Outcomes. 2014;12:176.

PARTICULARS OF CONTRIBUTORS:
1. Professor, Department of Cardiorespiratory Physiotherapy, KLE College of Physiotherapy, Hubballi, Karnataka, India.
2. Professor, Department of Musculoskeletal Physiotherapy, KLE College of Physiotherapy, Hubballi, Karnataka, India.
3. Assistant Professor, Department of Musculoskeletal Physiotherapy, KLE College of Physiotherapy, Hubballi, Karnataka, India.
4. Assistant Professor, Department of Neuro Physiotherapy, KLE College of Physiotherapy, Hubballi, Karnataka, India.
5. Assistant Professor, Department of Neuro Physiotherapy, KLE College of Physiotherapy, Hubballi, Karnataka, India.

NAME, ADDRESS, E-MAIL ID OF THE CORRESPONDING AUTHOR:
Prashant B Mukkannavar,
Professor, Department of Musculoskeletal Physiotherapy, KLE College of Physiotherapy, Hubballi, Karnataka, India.
E-mail: drshant24@gmail.com

AUTHOR DECLARATION:
• Financial or Other Competing Interests: None
• Was Ethics Committee Approval obtained for this study?   Yes
• Was informed consent obtained from the subjects involved in the study?   Yes
• For any images presented appropriate consent has been obtained from the subjects.   NA