

# Exploring the experiences of content experts with item vetting during item bank development

Anbreen Aziz<sup>1</sup>, Mashaal Sabqat<sup>2</sup>,  
Faiza Kiran<sup>3</sup>, Tayyeba Iftikhar Mirza<sup>4</sup>

## ABSTRACT

**Objective:** To explore content experts' experiences with item vetting during item bank development at a public sector medical university of Rawalpindi, Pakistan.

**Methods:** An exploratory study was carried out from December 2022 to February 2023 at a public sector medical college of Rawalpindi. A purposive sampling technique was employed to collect data from all content experts of the study institute who participated in item vetting activity during pre-exam moderation in the university. A pilot-tested semi-structured interview guide was utilized, interviews were audio recorded and later transcribed. Participants' anonymity was ensured. Various quality assurance strategies were employed to ensure the trustworthiness of the findings. Thematic analysis was performed on the transcribed data and themes were finalized by achieving consensus among all authors.

**Results:** Six themes overarching the fourteen subthemes emerged from the data. Participants expressed a profound sense of satisfaction and valued their experience in refining expertise in constructing multiple-choice questions (MCQs). It was widely acknowledged that such activities not only contribute to the enhancement of item development skills but also improve quality of items.

**Conclusions:** The consistent implementation of item vetting routines, in conjunction with diligent adherence to item writing protocols, contributes to quality assurance measures in assessment. Item bank development for fair and transparent assessment ensures production of competent healthcare professionals filtering incompetent ones hence improving health care services in the community.

**KEYWORDS:** Examination questions, Quality assessments, Medical education.

doi: <https://doi.org/10.12669/pjms.40.6.8664>

**How to cite this:** Aziz A, Sabqat M, Kiran F, Mirza TI. Exploring the experiences of content experts with item vetting during item bank development. *Pak J Med Sci.* 2024;40(6):1241-1246. doi: <https://doi.org/10.12669/pjms.40.6.8664>

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

1. Anbreen Aziz, BDS, MHPE, AFHEA  
Assistant Professor, Department of Medical Education,  
HBS Medical and Dental College, Islamabad.
2. Mashaal Sabqat, MBBS, MHPE  
Assistant Professor, Medical Education,  
Assistant Director Riphah Institute of Assessment,  
Riphah International University, Islamabad.
3. Faiza Kiran, MBBS, MHPE  
Assistant Professor, Department of Medical Education,  
Shifa School of Health Professions Education,  
Shifa Tameer e Millat University, Islamabad, Pakistan.
4. Tayyeba Iftikhar Mirza, MBBS, MHPE  
Assistant Professor, Department of Medical Education,  
Foundation University Medical College, Islamabad.

Correspondence:

Dr. Mashaal Sabqat, MBBS, MHPE.  
Assistant Professor, Department of Medical Education,  
Assistant Director, Riphah Institute of Assessment,  
Riphah International University,  
Al-Mizan Campus, Main Peshawar Road, Rawalpindi.  
Email: [mashaalsabqat@gmail.com](mailto:mashaalsabqat@gmail.com)

- \* Received for Publication: August 28, 2023
- \* 1st Revision Received: January 24, 2024
- \* 2nd Revision Received: February 22, 2024
- \* Final Revision Accepted: February 28, 2024

## INTRODUCTION

Valid assessment in a medical school ensures the production of competent doctors.<sup>1</sup> Knowledge, skills and attitudes gained by medical students are evaluated by a variety of assessment methods.<sup>1</sup> MCQs are written assessment methods frequently used to test higher order thinking skills.<sup>2</sup> They carry advantages of testing a wide range of content, assessing a large set of examinees and a rapid scoring system.<sup>2,3</sup> They are proved to be valid and reliable if constructed properly.<sup>2</sup> A good quality MCQ consists of a detailed clinical scenario-based statement, a clearly asked question in the form of a lead-in and one most appropriate answer along with three or four functional distractors.<sup>2</sup> Once an MCQ is constructed according to the recommended guidelines, it should be reviewed by the vetting committee which consists of an item author who must be a content expert, along with a medical education expert to review it from various aspects before testing in exams.<sup>1,4-6</sup> Vetting is a pre-test review and evaluation of items; items are corrected by detecting item writing flaws, checking content and

grammatical/language errors to improve their quality.<sup>4,7</sup> The psychometric properties of these vetted items are improved post-test, hence lead to development of item bank.<sup>4</sup>

Though literature is replete with studies signifying faculty training and item vetting in improving the quality of MCQs, many medical institutes do not employ item vetting strategy. Moreover, in Pakistani context, limited studies are available on the impact of item vetting towards improvement in the quality of MCQs. One of the reasons could be that the item vetting is not being carried out in most of the medical institutes of Pakistan. Therefore, the items used within medical colleges are often poorly constructed having flaws in them.<sup>2</sup>

A previous study from Pakistan exploring the barriers and facilitators in writing good quality MCQs recommends obtaining faculty perceptions after conducting workshops on item construction and performing item vetting.<sup>8</sup> Therefore, our study aimed to explore the perceptions of faculty regarding their item vetting experiences. This exploration holds significance in our context as it serves as a needs assessment for designing related faculty development initiatives, based on faculty input. Such thoughtfully designed faculty development programs will help enhance faculty competence in producing valid and reliable items.<sup>9</sup> This in turn, will lead to a fair assessment of students' abilities, ultimately leading towards the production of competent physicians, and improving health care services in the community.

## METHODS

A qualitative exploratory study was conducted in a public sector medical college of Rawalpindi from December 2022 - February 2023.

**Ethical Approval:** It was obtained by the ethics review committee of Army Medical College (ERC/ ID/ 244, Dated: December 12, 2022).

**Participants and procedure:** Purposive sampling was used to gain an in-depth understanding of the participants' perspectives. All content experts (Assistant Professor and above) of the study institute who were involved in item vetting activity during pre-exam moderation were invited to participate in the study via an email comprising a brief introduction of the study. In-depth, semi-structured interviews ranging from 30 to 40 minutes were conducted and recorded by AA, MS and TIM on phone, till data saturation, which was achieved at the thirteenth participant. The interviews were transcribed by MS and TIM.

**Interview guide development:** Comprehensive literature review was done by AA and the interview guide was developed based on the derived themes by discussion among all authors. The interview questions were validated by two medical educationalists, and pilot-tested with two subject experts.

**Data Analysis:** As the data constituted less than 500 pages, manual thematic analysis was performed,<sup>10</sup> and themes and subthemes were extracted by highlighting

key features, similarities, and differences in opinions of participants, and generating collaborative insights amongst the authors.

**Quality assurance strategies:** The authors used the quality assurance criteria for naturalistic studies by Lincoln and Guba for this study.<sup>11</sup> The strategies used for quality assurance included three peer debriefing sessions,<sup>12</sup> maintaining a reflective journal to note down the non-verbal cues of the participants during interviews,<sup>13</sup> and sending the interview transcripts to the respective participants for member checking to ensure credibility.<sup>15</sup> Moreover, data analysis was done by two authors, and the analysis was verified with all authors to ensure data analysis triangulation.

## RESULTS

A total of thirteen faculty members participated in the study: seven from basic sciences (53.84%) and six from clinical sciences (46.15%). These participants took part in a two-month item vetting activity, held in a public sector medical university from mid-July to mid-September 2022. This activity was held in the form of teams and each team comprised of three subject experts and one medical education expert from the affiliated colleges. The vetted items were planned to be part of the final professional examination of the university. The university granted remuneration to the team members. Data analysis generated fourteen sub-themes that were conjured into six themes. (Table-I)

**Organized and Fruitful Activity with Meaningful Learning Experience:** All participants found the opportunity to interact with subject experts and medical educationalists from diverse institutions, hence had valuable learning experience. They appreciated the systematic and organized process and expressed that the knowledge gained during these days was considerably more practical and useful. Many participants reported that they have learned about technical and grammatical errors, others learned about cover-uncover test and alignment between lead-in and options. In particular, basic sciences faculty noted a significant improvement in their ability to create application-level MCQs, which they previously found challenging.

**Impact on Quality Assurance in Assessment:** According to the participants, the activity increased the number of reviewed items, facilitating the development of item bank. The activity also improved content validity by segregating the items according to the table of specifications and highlighted many important learning objectives that were not being assessed previously.

**Recommendations:** The participants provided several recommendations for improving the item vetting process. They suggested giving due credit to the faculty members in the form of remuneration to motivate them, planning strategically, informing the faculty members in advance to enable them to manage their time well, and distributing the workload uniformly among faculty members to prevent exhaustion. Finally, the participants suggested a stepwise approach to item vetting: initial

Table-I: Themes, sub-themes and corresponding quotes generated by data analysis.

<i>Subthemes</i>	<i>Codes</i>	<i>Participant's quotations</i>
<b><i>Theme-1: Organized and Fruitful Activity</i></b>		
Systematic process	Systematic with both departments and university involvement Methodical & organized process	It is a good activity... because it is running now systematically at departmental as well as university level. Previously things were going haphazardly without any feedback and changes at university level. In my department I did it for many years, but now the whole process is more organized. (R5)
MCQ standardization	Useful due to standardization	A criterion has been set and everyone has to follow that. This helps in standardizing the MCQs. Everyone has to make MCQs of the same standard. (R1)
<b><i>Theme-2: Meaningful Experiential Learning</i></b>		
Hands-on/ Practical experience	Hands-on activity Practical skills-based activity	I learned by doing hands-on activity and gained skill of moderating MCQs; previously it was theoretical with no practical experience. (R15)
Input of multiple minds (Diverse faculty with educationalist)	Useful because diverse faculty and educationalists involved Interaction with other institutions' faculty	When you have more people in a team, they point out some sort of confusion in MCQ... then one realizes that students can also think along the same lines, then you modify the question accordingly. This has helped me a lot in making the stem of the MCQ crisper and clearer. (R7)
Multiple subject experts to remove bias	Multiple subject experts to remove bias Experts coming together Expert consensus	It was a very useful activity because, at times, we are unable to see the flaws while constructing the MCQ but when we look at it from the perspective of multiple minds, things get better. (R2)
Interaction with other college's faculty	Interaction with faculty from other institutions	I like the overall activity for engaging multiple specialists from different medical colleges. This helps to remove bias. (R10)
<b><i>Theme-3: Increased Confidence</i></b>		
Awareness about item writing flaws	Learned about alignment between stem & lead-in  Learned about technical mistakes in stem, lead-in and options Learned about grammatical errors  Learned about cover uncover test	They invited faculty of top minds from different institutes affiliated with this university. I was happy that we will be on the same page from now onwards. One regains confidence by interacting with faculty from other institutes (R11)
<b><i>Theme-4: Enhanced Test/Item Writing Skills</i></b>		
		Many times, through the help of the vetting team, I was able to replace the nonfunctional distractors with better options, which never came to my mind while making the MCQs. (R2)
		All types of flaws related to the stem and in scenario making, and asking lead in the form of question are improved. (R5)
		Most of the mistakes in my MCQs were grammatical errors. I found out that the stem and lead-in should be aligned. I remember we used to put the stem for the sake of having a scenario on the top; it had no relation with the lead-in. The options were never aligned; we would just put the options of our own choice. (R4)
		The cover uncover test was very new for me. I had never thought that each component of MCQ should be aligned. (R1)

<i>Subthemes</i>	<i>Codes</i>	<i>Participant's quotations</i>
Development of application-level items for basic sciences	Difficult to make application level MCQs for basic science subjects	It is hard to make application level MCQs/ scenarios in basic sciences subjects when curriculum is traditional, and students are not taught with clinical relevance. (R13)
	Learned about difficulty level of MCQs	Most common flaws in MCQs were ... too easy options which lower the difficulty index of item. A little hard work can improve the difficulty index and hence the quality of MCQ. (R14)
Applied the knowledge gained during certificate of health professions education (CHPE)	Putting CHPE knowledge to practice	I had learned the theoretical part of MCQ construction in my CHPE course, and when I tried applying that knowledge even then I was making mistakes.
	Skills were polished due to practice	When you have hands-on experience, you learn many things. (R3)
	Practical experience on top of basic knowledge	
<b><i>Theme-5: Impact on Quality Assurance in Assessment</i></b>		
Improved quantity and quality of MCQs	Improved quality of bank	The quantity of good constructed MCQs definitely has been increased after this activity. (R5)
	Improved quantity of MCQs	To become supervisor, I need to make many MCQs, and I think this exercise has helped me in learning how to make them. (R14)
Improved psychometric properties i.e., improved clarity of MCQs	Useful due to item bank development	My skills are polished due to this item bank activity. I can now make good quality MCQs. I believe practice and repetition of the activity make a man perfect. (R7)
	Improved clarity of MCQs	These questions will be given in exams and for the students it will be easier to understand the questions.
	Improved cognitive level of MCQs	Previously, some questions had no clarity, options had double meanings, now they are crystal clear. (R15)
	Better reliability and validity of MCQs	Regarding item vetting activity ... lets us review, recheck and re-evaluate; overall it improves the questions for all types of learners. Obviously, you can make it more reliable and valid. (R2)
	Positive impact on content validity of exam	Most of the Table of Specification (TOS) was ignored previously while making paper. This activity will be helpful in future while making papers from vetted pool. (R15)
Educational impact of the activity		I believe when the vetted MCQs are given to the students, they will study deeply, as the majority are problem-solving, and students will have to read their topics thoroughly in order to attempt these MCQs. (R3)
	Deep learning promoted due to improved quality MCQs	Of course, the assessment quality will improve. Students will have a deep approach rather than rote learning. (R8)
	Positive impact on student learning due to application level MCQs	...students will learn extensively as paper will come from all topics. They will also study applied aspects of basic sciences as clinical scenarios are added now. (R15)
<b><i>Theme-6: Faculty Recommendations</i></b>		

<i>Subthemes</i>	<i>Codes</i>	<i>Participant's quotations</i>
Fulfilling faculty needs	Due credit should be given to faculty involved.	Faculty who participates in this activity must be given an incentive as it is a time-consuming participation. (R5)
	Faculty members should be informed in advance.	The faculty members should be informed about the activity well in advance so that they can be prepared and also manage their time with their workload. (R14)
	Should be less strenuous for faculty	There must be a uniform distribution of workload by involving other members of the department, provided they are trained for this purpose. (R5) There should be a certain count to be vetted in a go, or in a setting because tired minds cannot give good performance. (R6)
	Workshop on MCQs and post hoc analysis needed	There should be faculty development workshops in which they should be taught, how to construct MCQs and then the faculty should be given the task to bring the MCQs that they have constructed or modified after the workshop so that they can be vetted by experts. (R9) I think I want to attend workshops on MCQs and post-hoc analysis to improve myself. (R13) There should be interdepartmental meetings which should discuss a few MCQs weekly, or bimonthly but regularly. After those meetings, you can take the vetted MCQs to the medical educationist to further assess the different aspects of a good MCQ. Finally, they can be vetted again in a centralized manner within university. Then it can be a more fruitful and productive activity. (R2)
Frequency of item vetting activity	Should be regular but at intervals to accommodate faculty workload (Frequency: once a year Frequency: at least twice a year Frequency: 4 times a year) Progressive item vetting at different levels	It should be carried out twice a year if not more than that. (R3) Whatever I have learned through this activity. Such activities should be held on a regular basis and each faculty of the department should be involved in such activities. (R8)

\*R= respondent.

scrutiny and improvement of items in departmental meetings, then assessment by medical educationalist, and finally, a central vetting process for enhanced efficiency and effectiveness.

### DISCUSSION

The aim of this study was to explore the perspectives of faculty members involved in the item vetting process at a university. We found that a well-organized item vetting activity has a positive impact on both the quality

of MCQs and the learning of the faculty regarding MCQ construction and its flaws. The common item flaws observed in this study were technical, such as longest option being correct, convergence, and negatively phrased lead-in. Furthermore, unfocused stems, and the use of excessive verbiage were noted in item construction, as also discussed in Tarrant and Ware's study.<sup>14</sup> Downing has explained that such errors are made mostly by novice item writers who are aware of item writing guidelines but fail to adhere to them.<sup>15</sup>

The current study confirmed Downing's reasoning as it was found that despite their prior participation in several workshops on MCQ construction, the MCQs submitted by faculty had multiple item flaws. The reason for the lack of a discernable impact of the aforementioned workshops could be a lack of experiential learning, as shown by the current study's results. Faculty members expressed appreciation for experiential learning facilitated by educationalists during the process. This aligns with the findings of Kiran<sup>16</sup> and Notzer<sup>17</sup>, who emphasized the transformational nature of learning during hands-on practice and feedback in skill enhancement.

Additionally, faculty members acknowledged their increased confidence in MCQ construction after this activity, as also highlighted in a study conducted by Al Faris et al.<sup>18</sup> Considering the two-month duration of the activity and its discernable impact on the participant's perceptions, our results mirrored those of a study by Iramaneerat, confirming that sessions on MCQ writing and analysis, regardless of their duration, enhance faculty members' test-writing skills.<sup>19</sup> We propose the implementation of more frequent, well-planned, and organized item vetting activities involving both junior and senior faculty members. Although this activity was demanding in terms of time, commitment, availability, and transportation logistics, faculty members found it more favorable as it provided a unique, albeit exhausting, yet meaningful learning experience.

This study has shown that item vetting in medical education improves item writing skills and satisfaction of faculty, transforms their perspective on item construction, improves content validity, and enhances quality assurance in assessment. As assessment drives learning, high quality MCQs assessing application of knowledge will impact students' learning style and promote deep learning in medical students. This, eventually, will have a positive effect on healthcare system by producing competent healthcare professionals, and filtering incompetent ones by robust assessment.

**Strengths of the study:** This study has highlighted the significance of incorporating quality assurance procedures in assessment.

**Limitations:** The subject experts from all affiliated medical colleges should have been included in the study.

## CONCLUSION

Regular item vetting exercises, accompanied by adherence to item writing guidelines, are widely embraced by faculty members due to their efficacy in yielding positive learning outcomes and improving content validity of assessment. This is primarily attributed to the inclusion of hands-on practice and timely provision of constructive feedback within the vetting process. Recommendations provided by the faculty should be helpful in future activities. Moreover, other universities can follow the same pattern of item bank development for fair and transparent assessment.

**Conflict of Interest:** None.

**Grant Support & Financial Disclosures:** None.

## REFERENCES

1. Salam A, Rabeya Y, Abu Bakar SM. Review article: Multiple Choice Questions in Medical Education: How to Construct High Quality Questions. *Int J Hum Heal Sci.* 2020;04(02):79-88. doi: 10.31344/ijhhs.v4i2.180
2. AlMahmoud T, Elzubeir MA, Shaban S, Branicki F. An enhancement-focused framework for developing high quality single best answer multiple choice questions. *Educ Health Change Learn Pract.* 2015;28(3):194-200. doi: 10.4103/1357-6283.178604
3. Sadaf S, Khan S, Ali SK. Tips for developing a valid and reliable bank of multiple-choice questions (MCQs). *Educ Heal Chang Learn Pract.* 2012;25(3):195-197. doi: 10.4103/1357-6283.109786
4. Gopalakrishnan S, Udayshankar PM. Question vetting: The process to ensure quality in assessment of medical students. *J Clin Diagnostic Res.* 2014;8(9):XM01-XM03. doi: 10.7860/JCDR/2014/9914.4793
5. Baig M, Ali SK, Ali S, Huda N. Evaluation of multiple choice and short essay question items in basic medical sciences. *Pak J Med Sci.* 2014;30(1):3-6. doi: 10.12669/pjms.301.4458
6. Puthiaparampil T, Rahman MM, Gudum HR, Brohi IB, Lim IF, Saimon R. How to grade items for a question bank and rank tests based on student performance. *MedEdPublish.* 2020;9(1):1-9.
7. Hassan S, Simbak NB, Yusoff HM. Structured vetting procedure of examination questions in medical education in Faculty of Medicine at Universiti Sultan Zainal Abidin Malaysia. *Malaysian J Public Health Med.* 2016;16(3):29-37.
8. Aziz A, Jaffer M, Qaiser A, Siddique L, Ali S. Barriers and Facilitators to Write Good Quality MCQs for Dental Assessments: A Qualitative Case Study. *J Ayub Med Coll.* 2022;34(1):178-182.
9. Karthikeyan S, O'Connor E, Hu W. Barriers and facilitators to writing quality items for medical school assessments - A scoping review. *BMC Med Educ.* 2019;19(1):14-17. doi: 10.1186/s12909-019-1544-48
10. Creswell JW. *Educational research: Planning, conducting, and evaluating quantitative and qualitative research.* Fourth Edition Published by Pearson Educ. 2012.
11. Egon G. Guba YSL. Epistemological and Methodological Basis of Naturalistic Inquiry. *Educ Communit Technol J.* 1982;30(4):233-252.
12. Sabqat M, Khan RA, Yasmin R, Mahboob U. Exploring developmental assignments shaping experience-driven acquisition of leadership competencies in young clinicians. *BMC Med Educ.* 2022;22(1):1-11. doi: 10.1186/s12909-022-03544-y
13. Johnson JL, Adkins D, Chauvin S. A review of the quality indicators of rigor in qualitative research. *Am J Pharm Educ.* 2020;84(1):138-146. doi: 10.5688/ajpe7120
14. Tarrant M, Ware J. Impact of item-writing flaws in multiple-choice questions on student achievement in high-stakes nursing assessments. *Med Educ.* 2008;42(2):198-206.
15. Downing SM, Haladyna TM. *Handbook of test development.* Lawrence Erlbaum Assoc. 2006.
16. Kiran F, Ayub R, Rauf A, Qamar K. Evaluating the impact of faculty development programme initiative: Are we really improving skills in MCQ writing? *J Pak Med Assoc.* 2021;71(10):2434-2438.
17. Notzer N, Abramovitz R. Can brief workshops improve clinical instruction? *Med Educ.* 2008;42(2):152-156.
18. Al Faris E, Naeem N, Irfan F, Qureshi R, Saad H, Abdulghani HM, et al. A one-day dental faculty workshop in writing multiple-choice questions: an impact evaluation. *J Dent Educ.* 2015;79(11):1305-1313.
19. Iramaneerat C. The Impact of Item Writer Training on Item Statistics of Multiple-Choice Items for Medical Student Examination. *Siriraj Med J.* 2012;64(6):178-182.

## Authors Contribution:

**AA:** Conceived the idea and made a proposal.

**MS and TIM:** Collected the data.

**AA, MS, FK and TIM:** Did data analysis and interpretation.

**AA, MS and FK:** Drafted manuscript.

All authors finalized the manuscript before submission and are accountable for the integrity of research.