ELIMINATING FLAWED ITEMS IN HIGH-STAKES EXAMINATIONS
THE CHINESE UNIVERSITY OF HONG KONG EXPERIENCE

KUMTA Shekhar M., **JIN Yan**, LEUNG Y. C. Joseph, YUNG L. K. Alex

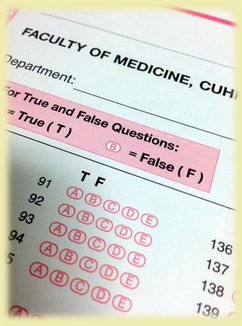
Teaching and Learning Resource Centre, Faculty of Medicine, The Chinese University of Hong Kong

香港中文大學 The Chinese University of Hong Kong

香港中文大學醫學院 Faculty of Medicine The Chinese University of Hong Kong

Teaching & Learning Resource Centre

Background Information



Quality control and quality assurance of examination questions for assessing and certifying medical student achievement is crucial. It is the responsibility of institution to seek ways to ensure quality of high-stakes examinations.

Teaching and Learning Resource Centre (TLRC) of The Chinese University of Hong Kong (CUHK) completed a major revamp of items used in its final surgical examination in Year 2013.

An internal review of examination item completed in 2007 revealed that half of the items (311 out of 625) that were used in final surgical examination from 2002-2005 were flawed[1].

Table: Flawed and non-flawed items divided by cognitive level tested

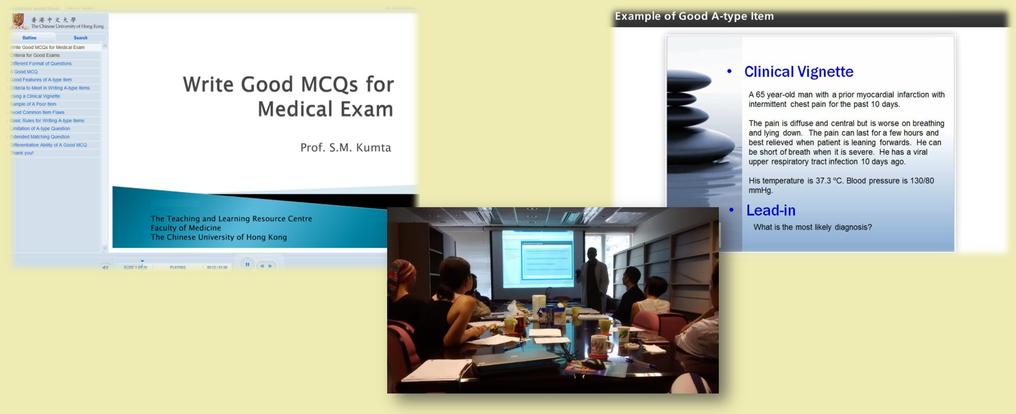
(Cohort I 2002-2005) n= 625

Cognitive level	Flawed	Non-flawed	TOTAL
K1(recall)	287	223	**510**
K2 (reasoning)	24	91	**115**
TOTAL	**311**	**314**	**625**

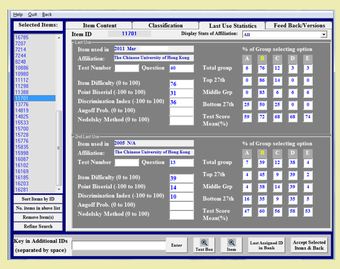
TLRC attempted to remediate the situation through providing bi-annual item writing workshops, reorganization examination committees, and most importantly, implementation of an item bank from which items for assessments can be store and viewed.

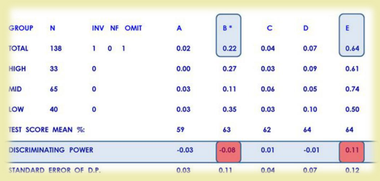
Summary of Work

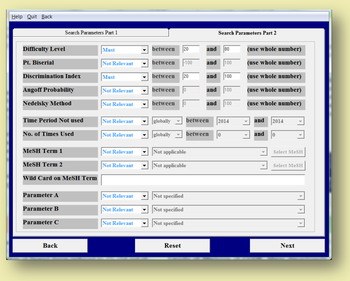
The item writing workshops held by TLRC teach teachers with guidelines[1] on how to create quality questions and the examination committees screen all questions that are used in high-stakes examination.



An item bank called IDEAL[2], i.e. International Database for Enhanced Assessments and Learning (IDEAL), is widely promoted to departments for storing all the items used for high-stakes examinations. IDEAL allows each examination question and its statistical data (after exam) to be stored and collected after examination. It provides specific indicators of potential flawed items that need further review.







Summary of Result

A follow-up study on the items used in the final surgical examination for 2008-2011 was performed and number of flawed items reduced to 7% (54 out of 742).

Table: Flawed and non-flawed items divided by cognitive level tested

(Cohort II 2008-2011) n=742

Cognitive level	Flawed	Non-flawed	TOTAL
K1	26	374	**400**
K2	28	314	**342**
TOTAL	54	688	742

Description of flaw	Cohort I (2002-2005)	Cohort II (2008-2011)
• Grammatical clue in sentence completion	1	1
• Logical clues in stem relating to best option	0	1
• Word repeats in stem and best option	0	0
• Convergence in best option	1	3
• Longest option is answer key	45	8
• Format failure, lost sequence for data	4	0
• Use of absolute terms in options	1	2
• Use of vague terms in options	6	5
• Use of negative term in stem	95	7
• "None/All of the above" as answer key	16	2
• Failed Hand Cover Test, or unfocused stem	139	9
• Unclear English and or ambiguities	3	16
Flawed questions vs Total	311 vs 625	54 vs 742
Percentage	50%	7%

Repeated exam items are dramatically reduced from Year 2008 to 2011. Psychometrical data stored with individual item in IDEAL item bank helped on item quality assurance.

Conclusion

TLRC interventions greatly improved the quality of the items that are used in high-stakes examinations in the Faculty of Medicine of CUHK.

More departments in the faculty adopted IDEAL as the database to store their own examination items. Item analysis is regularly done after exam to ensure the quality assessment of high-stakes examinations.

Take-home Message

- The use of an item bank with item analysis program is extremely useful and important for identifying problematic items before and after examination. It ensures the examinations quality and provides a qualified assessment for students.
- Training on item writing can help teachers to write quality questions for medical examinations

What to Expect

- The IDEAL item bank is under revamp to make it more user friendly. It may help users to improve quality items more easily.

Reference:

1. Item Writing Guidelines: Case SM., Swanson DB. Constructing written test questions for the basic and clinical sciences. NBME, Philadelphia. 2001.
2. International Database for Enhanced Assessments and Learning http://www.idealmed.org/homeindex.html

Contact: tlrc-elearning@med.cuhk.edu.hk