

Review of the Part 1 MFPH examination: Autumn 2005

Dame Lesley Southgate

3/2/2006

Introduction

The review of the MFPH Part 1 examination has been conducted against a background of unprecedented change in postgraduate training in medicine, and a new climate for external quality assurance of all aspects of that training, to be determined by the Postgraduate Medical Education and training Board (PMETB). This has led to a wider review than would have resulted from an enquiry solely into the events of May 2005 that prompted the Faculty to request this report. The Faculty is to be congratulated on taking this initiative before introducing changes. This report should be read in context: there are many people committed to delivering an excellent examination, and who work hard to that end. But there is also wide acceptance that there is room for improvement, with the recognition that the work of the examiners and the examination department is under-resourced. The Faculty will have to make some tough choices at the highest level about providing adequate resources if it accepts the recommendations I have made.

Background

PMETB QA and MMC

Since its establishment, the PMETB has published a series of papers on curricula for specialty training in medicine, principles for assessment programmes, internal quality control for assessment systems, and on workplace assessment.¹ All of these papers are important background for this review. Together they set out the direction of travel for work that must be done to ensure that the MFPH is part of an assessment programme that enables the public, the profession and the candidates/trainees to have confidence in those decisions taken on the basis of examination results. The examination provides only one part of the evidence for determining the progression of the trainee, and passing or failing is considered alongside the information from in-training assessments in the workplace. However, if failing the examination can inhibit or stop progression, it becomes a very high stakes assessment indeed, and the spotlight will inevitably move to the quality of the examination and the reliance that can be put on its outcomes. And if there is significant doubt about the exam, the reputation of the Faculty will also be in the spotlight.

The first priority for the work ahead will be to understand the implications of the PMETB papers for the integration of assessments with the curriculum approved by the Board. Defining the primary purpose of Part 1 as part of the public health

¹ All of these papers can be downloaded from the PMETB website.

assessment programme, and appreciating the effect of this on the selection of content and the timing of the examination, comes next. And then there will need to be decisions about the limitations of the format and what is better assessed in the workplace, or by different methods. These considerations may affect committee structure within the Faculty, and committee membership, including the participation of members of the public and trainees. Accordingly, I have placed considerable emphasis on them in this report. The impact of the multi-professional and diverse cohort who take the examination has also been considered and its impact on the purpose and content of the examination.

Events in May 2005

On 6 May 2005 many candidates taking the Part 1, Paper 11B, encountered difficulty in understanding one of the questions on the paper. This led to an escalating situation reflecting the distress of some of the candidates and revealed flaws in the arrangements in place to deal with such a situation despite the best efforts of the staff who were present. The events have been detailed both in the report from the Head of Education and Training and from the report submitted to me by the members of the Trainee Members Committee. Both accounts relate the same events from different perspectives but are not in conflict with each other. The Trainee Members Committee have submitted a careful and well researched account which I have taken into full consideration in coming to my conclusions in this report. ***I have not appended either document, but I suggest that they be made available to those who are charged with considering my conclusions.***

On reading both reports about that day, it is clear to me that these events reflected deeper problems with test construction, security and lines of responsibility for final decisions about the content of the paper. All of these, and other, issues will be dealt with systematically and linked to the PMETB Principles for Assessment and discussed under the findings and recommendations.

On 6th May a critical incident occurred, an 'accident waiting to happen'. Like all such incidents, there were many contributory factors, and I have therefore avoided apportioning blame or naming individuals. That would be too easy, too simple, and counterproductive. If the recommendations at the end of the report are considered and implemented, repetition of this situation will be avoided because in some sense, problems with each of those areas contributed to what occurred.

Data collection

The data, which form the basis for this report, were collected by means of interviews, perusal of all documentation available for examiners and candidates, internal papers summarising operating procedures for the conduct of the examinations, and inspection of the test banks. I also read the papers from the Part 1 May 2005 sitting and can confirm that I too, found the trainees complaints about the clarity of questions 2 and 4 to be justified.

Interviews

Face-to-face, in-depth interviews were conducted by DLS with eight individuals who have, or have recently had, a leadership role in the examination or the faculty. Telephone interviews were also conducted with examiners and trainees by researchers from the OUCEM². These interviews were based on a strict written protocol derived from the PMETB Principles for Assessment.

Findings

Findings are presented under headings derived from the PMETB principles for assessment. They are based on a synthesis of all of the evidence considered. Illustrative quotes, shown in italics, are taken from the face-to-face interviews conducted by DLS. The full report of the telephone interviews is appended.

The purpose of the examination

Clarity of purpose of an examination is fundamental to all of the other attributes that it must exhibit. All of the interviewees were asked about the purpose of Part 1 MFPH and there was a consensus about the primary purpose.

Quote: Typical response

- *The Part 1 tests the knowledge base and the application of knowledge. It is intended for the earlier part of training and represents the knowledge that is necessary for the trainee to move from learning to action.*
- *It implies that they take more responsibility and that the standard is set to ensure public and patient safety.*
- *We would expect a pass rate for trainees taking it for the first time to be about 70%. The timing of their attempts makes a difference, and selection is a problem.*

There is wide agreement that the Part 1 examination is intended for trainees who must demonstrate that they have the knowledge base that will act as the foundation for all future public health practice. It is also clear that both medical and non-medical trainees must have the same knowledge base at this early stage of training and that the examination is intended for the whole cohort. There was little mention of issues to do with devolution. The examination is intended for all UK trainees.

The examination provides part of the evidence to support the RITA process and persistent failure will contribute to a RITA E. This implies that the examination establishes the minimum standard for the knowledge base that must be achieved to progress in specialist training and ultimately to a CCT for the medics. But it also implies that it is part of a wider assessment programme, much of which is based in the workplace, to support the RITA process. This has implications for the content of the examination, co-ordination with other in-training assessments and the passing standard that is set.

² The Open University Centre for Education in Medicine.

During the interviews, and from documentation, it became clear that many other candidates who take the Part 1 exam do so for different purposes than those outlined above.

Quote

- *Some people use it for CPD.³*
- *Others may use it for Article 14 applications in the future.*
- *The examination is taken by a mixed cohort but during the early years the syllabus is the same for medics and non-medics.*
- *Employers should be able to say that anybody who has taken this exam should have a defined knowledge base with the ability to apply it.*

Some established public health practitioners may take the examination to demonstrate that their knowledge base is up-to-date, others may take it to enhance their evidence for application for direct entry to the Specialist Register under Article 14 of the legislation which has established the PMETB. Many of these candidates may practice in very different healthcare systems and deal with public health problems within populations where the epidemiology of disease is very different. If the examination is to serve the primary purpose above, the implications for test content, difficulty level and fitness for purpose for these other groups are profound. It will be difficult to design an examination that can meet all of these purposes, and it is likely that the present test, in seeking to do so, leads to pass rates which are very different for different cohorts. And a knowledge base for the early part of training may not be relevant for CPD purposes.

The content of the examination

The validity of an examination is linked to its content. If the content is not well defined and if it does not sample the domain adequately, the test will be flawed. The purpose of the test is also fundamental to ensuring appropriate content. While there is a syllabus for the Part 1 MFPH examination, it comprises a list of topics without adequate specification of the level or complexity of knowledge and understanding that the candidate should be able to demonstrate. It was reported to me that recent work to map the syllabus to the world of public health practice has been carried out by the examiners in order to ensure coverage of relevant topics. However, it was reported to me by one interviewee that this work has been in isolation from those in the Faculty who are developing the curriculum for training. This is a major barrier to establishing validity and fitness for purpose.

Typical Quotes

- *The examination is referenced to the syllabus. It is a long list and needs improving to reflect current practice. It needs to be referenced to the world of work and the world of training.*

³ This is a direct quote from a senior member of the Faculty. But in feedback on the draft of this report, it has been mentioned that established members will not take it for this purpose. The general point stands.

- *The long list was circulated to examiners, all of the main MPH programmes, regional advisers and universities teaching public health.*
- *The fit with the curricula for the MPHs around the country are different, and the trainees know this. They probably provide about 70% coverage.*
- *The RITA competencies and the exam syllabus contain the content. There are Masters programmes run by universities that programme directors pay for their trainees to attend. There may well be uneven quality between these programmes, and hence uneven preparation for the exam. How much all of them prepare trainees for the content of the examination might vary. The programmes are not always principally designed for medical trainees, and some are taken by international students for practise in their own settings.*
- *The curriculum does not exist at present. There is a syllabus for Part 1 and four competencies for Part 2 which are presented as headings with lists of activities included to guide the candidate in preparation.*
- *The RITA competencies framework is important*
- *There is no formal blueprint to guide paper selection for part one*

And the trainees have highlighted this area as a major concern.

- *One of the concerns about the syllabus for the examination is the lack of clarity of what is expected of the candidate. There are examples of topic lists that are not further clarified by specific topics; for example, elementary human genetics.*
- *At present, the feedback on examination questions is not linked to components of the syllabus. The 'key points' feedback system in place is variable in its comprehensiveness. It also has little reference to the syllabus. Each question for Part 1 should have a list of components of the syllabus that are being assessed.*

The PMETB has issued standards for curricula, which will be applied when approval is requested. And the Principles for Assessment Systems are designed to work together with those standards to make a coherent framework for Colleges, Faculties and Deaneries to support the trainees in achieving a CCT.

The Faculty must develop its curriculum for training in accordance with those standards, and that means that the content will be referenced to public health practice, but that the level that trainees are expected to reach, and the knowledge and skill required for each stage of training is explicit and clearly set out. **The examination must then sample that content, at the level determined by the purpose of the examination.**

Hence the Part 1 MFPH will be sampling the knowledge and applied knowledge set out in the curriculum for the early part of training, which will include a specification of the complexity and depth expected at that stage. A test blueprint should therefore be constructed, based on the curriculum, and taking account of where content is best assessed: in Part 1 or in Part 2. This is fundamental to establishing the reliability and validity of the test. Other aspects of the curriculum will be assessed in the workplace, or through participation in the MPH programmes. Trainees will be able to judge the relevance of these programmes to their education and training only when they see a detailed and well-constructed curriculum set out for their consideration. However, the Part 1

examination is but one element in the assessment of training, and if it is to be deemed appropriate to make it mandatory to pass, trainees, the public and employers have a right to expect that it is constructed to the above standards.

These remarks are aimed at making the examination fit for the primary purpose of assessing trainees in the training programmes. The Faculty opens the examination to many other candidates. They should be aware of the content of the examination from inspection of the curriculum and documentation about the examination. They can then decide if it is an examination that is relevant for their own purposes. And for its part, the Faculty should avoid the inclusion of content that is too context bound. Asking questions that have different answers in different parts of the country should not be part of this test. The situation is much more difficult for international graduates and a review of test material with their needs in mind will be another priority. But this difficult area must be dealt with as it introduces unfairness into the exam if it is not considered.

Assessment methods and constructing test material

Once the purpose and content for an examination is determined, the test method may be selected. The Part 1 is held to test the knowledge base, and it is appropriate that this should be done by a national, written test. However, the crisis in May has highlighted threats to validity and reliability that must be addressed. The procedures that are in place for governing test construction, writing test material, constructing marking schedules and lines of responsibility for the final content of the paper look adequate. However, it is clear that they are not strictly followed despite the effort of the Examination Department to ensure that they are.

The Faculty has constructed a test bank, but the material in it is principally headings or incomplete questions, rather than fully worked up questions with statistics about their facility and discrimination recorded for each time they are used. I found it difficult to find any data that are collected about the examination that would allow the difficulty of any paper to be established, and it seemed that at each sitting, papers comprised different questions with no equating between test forms to ensure a consistent standard. The whole area of selecting an appropriate form of written test material is beyond the scope of this review, but the Faculty must follow best practice in developing the bank. This should include finalising questions before the deadline and agreeing the model answers at the same time. Formats could range from multiple choice questions in the EMQ or SBA format, through to short answer questions with a marking schedule which reflects key features of a passing or excellent answer. They are all reasonable choices. If the answers are difficult to agree, or likely to change within a very short time-span, the validity of the question for this type of test is called into question.

The one absolute that was clear from the events of May was that questions were constructed where the answer to one depended on information or answers to a previous question. This is considered to be poor practice and should be absolutely avoided in future.

Quote/interview notes from one interview

- *A general discussion about item writing and construction of the test bank and paper selection, revealed that methods for the assembly of papers, poor security practices and failure to enforce deadlines and carefully proof read papers make the Paper 1 unlikely to be of the highest quality.*
- *The change from professional printing to in-house production looks unprofessional and trainees might feel that they are not receiving value for money if there is a 'homemade' feel to the exam.*
- *Marking schedules are not completely agreed before the paper is administered and there is an element of fixing things on the hoof which is high-risk behaviour.*
- *Ten short answer questions, in Paper 1, especially if the answer to one question may depend on correctly answering a previous one, may not be the ideal format for testing the knowledge base and alternative item formats might be explored.*
- *There is no systematic data collection to enable the performance of individual questions to be documented, including facility and discrimination.*

Test security and proofreading

One of the characteristics of the problems that beset the May paper was the failure to enforce deadlines to agree the content of the paper.

If this situation recurs, test security must be called into question, as it may compromise the fairness of the examination. And there appear to be no checks in place to detect cheating, although I may not have been made aware of them. Measures to stop the circulation of test material by email, especially to work addresses must be introduced and a secure means of communication put in place. Ideally test material should be finalised, complete with model answers, and in a secure bank from where it is selected centrally so that examiners have no knowledge of the content of the examination paper before it is administered. Under no circumstances should amendments be dictated over the phone or written by hand, to be interpreted by secretarial staff before inclusion in the paper. And the final proofreading and authorisation to print must be in the hands of a senior examiner who takes responsibility for its accuracy. All of these measures must be implemented at once if they are not already in place.

Examples of other difficulties caused by a lack of attention to detail at the proofreading stage, or provision of background material during the examination, have been mentioned by the trainees. They were confirmed by the Head of Education and in my view, on both occasions the reliability of the paper will have been affected by the errors. Mistakes like these make the examination appear unprofessional, and should never occur at this level with a high stakes assessment.

The trainee report says:

- *In Paper IIa of the June 2004 examination, question 1 asked candidates to write a structured abstract of the paper provided. In square brackets, there was a statement indicating that objectives were provided. However, no objectives were provided. The invigilator queried it with the chair of the examiners and the candidates were told to state their own objectives. This was ambiguous and caused confusion amongst candidates.*
- *In Paper IIb of the January 2002 examination, question 2 referenced to the Majeed model, but the referenced paper was not provided. The FPH admitted that the omission of the referenced paper 'did not allow candidates to make a reasonable attempt at this question' and did not allocate any marks to responses to this question. The re-allocation of marks to another question was arbitrary.*

Standard setting

If we accept that the primary purpose of the test is to establish a knowledge base for the whole cohort of trainees for the early part of training, and if we take into consideration that trainees can be prevented from progressing to CCT if they fail, (although I have no information about how many attempts they are allowed,) then it seems that the expectation is that about 65-70% of first-time takers in training programmes should pass.

The rules for marking are clearly set out, however there are no data that establish the difficulty of each paper relative to the next one, and as this is likely to be related to the selection of test content and the quality of test material, the actual standard for any paper is hard to determine.

There are several methods described in the literature for setting standards for knowledge tests.⁴ The pass rate for this examination should be high (80%) because it establishes the basic standard for trainees to progress. Modern content-based standard-setting methods could be used to set a standard based on this expectation. However the actual pass rate in such a test will vary according to several areas of uncertainty. In principle, trainees should all be capable of passing if selection and training programmes are conducted well. But, in the absence of a well-defined syllabus, properly sampled, which is explicit about content and difficulty at the early stage, it is difficult for trainees to know what to prepare for, and hence capable candidates might fail.

Equally, if Masters programmes do not address the syllabus, completion of those programmes will not necessarily be connected to performance in the examination. However, completion of a Masters programme may support the trainee to achieve other relevant objectives for progression to CCT.

A mixed cohort take this examination, and as the purpose is to establish the minimum standard for progression, as a start, to deal with the problems

⁴ This is a specialist area and the Faculty will need advice on which method to adopt. A full description of the approaches is beyond the scope of this report.

identified above, the standard should be set against criteria, not against the performance of the cohort. Many UK examinations at this level use recognised methods to set standards based on the content of the test, with the standard set by experienced judges. This approach can only be adopted if test material is ready, with answer keys finalised, before the examination is administered. Once the standard is established, and the standard error of measurement around it known, then meaningful studies about selection and training and their relationship to candidate performance (and they are always difficult to do) could be contemplated.

Recommendations

1. The Faculty should make a clear statement of the primary purpose of the Part 1 examination. It is a test of the early knowledge and skill base of trainees for the practice of public health in the UK. In so doing it should make clear the implications for other candidates who take it for different purposes and from different settings.
2. The Faculty should press on with writing the curriculum for UK training in accordance with PMETB standards. The Part 1 examination should assess content defined in the curriculum to the standard defined within the curriculum for the early stages of training.
3. The examination should be blueprinted against the curriculum within an overall blueprint which takes account of which aspects are best tested in a national examination, and which are best assessed in the workplace. (RITA competencies) MPH programmes should be considered within this framework and steps taken to ensure that programme directors around the country are aware of this approach.
4. There should be urgent consideration given to the test formats chosen for assessing knowledge. The construction of a test bank is a good beginning but the quality of the test material needs review. The Faculty should consider asking a consultant on writing examination questions to undertake a detailed review and to run some item writing workshops for examiners.
5. The test bank should contain completed questions with model answers periodically reviewed for accuracy and currency. Material should address areas that are stable, which underpin the practice of public health, and which are not susceptible to changes in healthcare policy.
6. Test security needs urgent review and immediate institution of good practice.
7. Enforcement of deadlines for test construction and finalising the content of papers is a priority. A senior examiner must take responsibility for proofreading the paper and signing it off for printing. This includes

ensuring that all ancillary material necessary for answering papers will be available for candidates on the day.

8. A senior examiner should be present at the examination.
9. The Faculty should take steps to adopt criterion referenced (content based) standard-setting procedures as part of its commitment to the development of the examination in line with PMETB QA principles.
10. The resource for the Education Department, and support for senior examiners must be reviewed. There should be no reliance on untrained or temporary staff to undertake tasks where high level judgement or knowledge of public health is required.

Appendix

REVIEW OF FPH PART 1 EXAMINATION REPORT OF PRELIMINARY INTERVIEWS
--

Method

Semi-structured interviews were carried out using a pre-determined interview schedule (Appendix 1).

All interviews were conducted by telephone. Interviews were transcribed and subjected to content analysis by three individual researchers.

A total of 12 telephone interviews were carried out on examiners, administrators and trainees.

The sample consisted of:

Category of Interviewee	Number
Part 1 Examiner	3
Part 2 Examiner	3
Administrator	4
Trainee	2

Results

The purpose of the examination

The following table shows that most examiners and administrators are agreed that the exam is there to ensure that candidates have sufficient knowledge, skills and understanding, or to assess competence. Respondents did not comment on the role of the exam in relation to the RITA process tests or its structural role in training as a whole, or its relationship to the curriculum.

Comments: examiners and administrators	Number
To ensure there is sufficient breadth/depth of knowledge, understanding, skills, and attitudes in public health.	8
To assess competence/set a standard/make sure people are competent	5

The two trainees had similar views. But one, unsurprisingly, cited the need to pass the exam to progress in training:

- To analyse knowledge in public health;
- To test knowledge and skills in communication and numeracy;
- A hurdle to jump to the next stage.

Does the exam meet its purpose?

Respondents were divided as to whether the examination met its purpose. Six were sure that it did:

It is a good discriminator of people who do not have the minimum knowledge and understanding to pass safely through to competency exams.

However, there was concern from one respondent that it is not blueprinted against the curriculum:

It is not a modern reproducible exam in the American sense, but we are a small Faculty and the case has not been made to the members that the increase in fees required is worthwhile given the small risk that it would be proven to be not fit for purpose.

On the other hand, one commented that:

It is clear what they are looking at: a range of knowledge and some skills.

Who are the intended candidates?

There was no clear consensus about this. Three respondents felt that the exam was for ‘*people on SpR training schemes or a multidisciplinary training scheme.*’ Two felt that it was for ‘*people interested in public health but not on the training programme.*’ A variety of other suggestions were made.

Perhaps these responses suggest that there is insufficient information about the role of the exams in the training structure.

Is the exam mapped to a specific curriculum?

Although two felt that the exam is based on the Faculty syllabus, in general, respondents’ answers suggest some confusion and that this is a major area for development and clarity:

It’s linked to the syllabus and anything else the examiners thing suitable.

Not clear how it relates to the curriculum.

There are four competencies for Part 1 which are presented as headings with a list of activities to guide the candidate in preparation.

No overall blueprint for the exam.

Different universities might run different courses and might not be consistent in preparing trainees for the exam.

Trainees pointed out that;

The syllabus is on the FPH website, but the exam does not match it.

The syllabus is too broad for the exam.

They need to decide what they want us to know and how much we should know.

The evidence base behind the chosen components of the exam

Respondents were uncertain about the evidence base underpinning the examinations. While three thought that there was no evidence base, other views were expressed:

It's competency based in relation to many years of experience of what public health people ought to know to be able to do their job properly.

It's done pragmatically because of [lack of] resources.

It is not evidence based, really. But it is extremely unlikely that the knowledge and skills will not enhance becoming a public health consultant.

Are the other assessment methods that could be used?

There was no consensus view about other possible assessment methods. Three respondents suggested that workplace-based methods might be considered, though doubts were expressed about their validity and reliability. And inevitably, the question of lack of time to develop new methods was mentioned.

Neither trainee interviewed was happy with the current test methods. They worried about validity and felt that:

They are not measuring competency.

More exams, longer OSPHEs and structured scenarios might, for them, be better.

Is the exam content appropriate to the candidates' level of experience?

The two trainees disagreed on this point.

Is the current system reliable, valid, feasible and cost effective?

Again, no consensus view emerged. But, overall, there was a feeling that it is probably acceptable even though these technical characteristics of the exam had not been formally and sufficiently addressed:

It is an accurate indicator of the required breadth and depth of knowledge.

Results in different parts of the paper seem consistent. But some people do badly on some types of question. It has evolved and it is reviewed.

People worldwide answer the same questions at the same time.

People who pass the membership have, on the whole, been competent to practise.

I'm not sure that the same standards apply to England and Ireland. We need to set up co-operative arrangements

On the other hand, one trainee expressed a deeply felt concern:

I was shown the model answers for the questions I got wrong and the answers did not seem even related to the questions. I would never have got them correct.

Where the truth of this lies, we cannot say without further enquiry. However, any high stakes assessment, such as this, must be able to demonstrate to all its technical characteristics of validity and reliability.

In terms of cost effectiveness, again, respondents were not sure. No comparative studies have been done but it was suggested that item banks and optical marking might be researched.

Rationale and methodology for setting a performance standard or passing score

This seems to be another grey areas in the minds of respondents. Although two admitted that they did not know, others offered a variety of processes:

- *Whole paper for marginal candidates is reviewed and a decision made about whether they are fit to pass.*
- *Moderated in the candidate's favour.*
- *Pass rate is fixed at 50% and the answers are adjusted around this.*
- *More than one bad fail is an overall fail.*
- *We work it out at the end of every exam.*
- *Marking schedules are not completely agreed before the paper is administered and there is an element of fixing things on the hoof which is high-risk behaviour.*

Whatever the actual method is, it is possibly not made clear to all. None of the methods described would meet proper standards for setting pass marks. So perhaps both practice and information about this aspect of the exam are required.

Interaction between performance in difference parts of the exam

No respondent was clear about whether or not the performance in one part of the exam would affect performance in a subsequent part.

Security of test material

Although three respondents did not know anything about security of the item bank, others mentioned a variety of issues:

I was disturbed recently when they circulated comments on the wording of a new question to all examiners on their work emails.

Questions are available on the website, but get reused.

It depends on personal responsibility and security of post and emails.

Some elements of security were also mentioned:

This is taken care of by the Education Department.

Electronic communication is password protected.

Papers are shredded and comments sent by letter, not email.

This aspect of the examination seems to require some development and systematisation.

Feedback to trainees

Few respondents seemed to know what feedback candidates receive. Two examiners/administrators reported that candidates are given their marks only if they fail, while two others thought that all candidates receive their marks for each question. The two trainees knew that marks were only given to those who fail and they felt that there was insufficient feedback.

Pass rates.

Trainees believe that the pass rate is 20-22%. Not surprisingly, they felt that it as demoralising to fail.

Recruitment and induction of examiners

Most respondents reported on at least one aspect of selection and induction. If these are added together, they make the basis of a suitable system, but some development work would be required:

*Need to be experienced as a consultant.
Strict induction process - observed several exams. [3]
Attended classroom training session.
Compared my marks with other examiners.
Answered an ad.
Had to read a load of papers for induction. There was no induction process as such.*

Ongoing development of examiners

No-one reported any ongoing development activities for examiners.

Processes and standards for marking papers

Only two respondents claimed any knowledge in this area:

It is double marked according to model answers. Marks are averaged to review discrepancies between examiners. If the candidate is borderline and the examiners don't agree on a mark, then it has to look into closely.

How might the examination be improved?

There were many suggestions for improvement of the examination, but no consistent view:

*Quality assurance with Ireland.
More alignment with the syllabus.
Move to MCQs for some of the factual knowledge.
Someone has to take personal responsibility for question wording. The continual rewording of questions was what led to the disaster we had last year when a chunk of text was missing from a question that made it incomprehensible
Improve organisation and procedures of the exam.
Use a proper marking system to eliminate bias.
Review every exam at the end of it to learn from experience.
Make it to clear to examiners and trainees what is expected.
Blue print to the curriculum.
Need an expert on educational theory to look at it.
Give feedback to trainees.*

These are all sound suggestions that need to be put together in a quality improvement cycle. But the same concerns arise:

Who will pay for changes to the exam? Faculty members won't agree. We are under-resourced.

The two trainees also had many ideas:

Organise the exam. Decide what you want to test.

MCQs might be better.

Tell us what you re going to assess.

Improve the questions. Some of them are confusing. We spend too much time interpreting them. I disagreed with their answer to one question but there is no-one to tell that they are wrong.

Questions are vague.

It's hit and miss whether you answer in the way they're looking for.

You can be functioning well in your job and know your stuff, but this may not be the right approach for the examiner.

Resources

Although the issues of resources was not directly asked about, a number of respondents mentioned this as an issue and a barrier to developing and administering the examinations properly, as quoted above. This is clearly and issue for a small Faculty in the face of the need to develop the assessment system and its components to meet PMETB standards. Whatever, decisions are made, they must be taken in light of their resource implications. As one respondent said:

We are under resourced for all exam work and it is difficult to do things to a high standard.

Conclusions

This brief and limited study suggests that there are clearly causes for concern in the examination and its administration. But there is also a lot of thought and many ways forward have been identified. The underlying issue of funding for a reformed exam process is a recurring theme.

APPENDIX 1 INTERVIEW SCHEDULES

Trainee Questions

1. What do you think is the purpose of the exams?
2. Is the exam mapped to a specific curriculum? If so, which?
3. Are you happy with the current methods used to assess your competence? If yes, explain. If no, how do you think it could be improved?
4. Do you know what the pass rates are for each section of the exams? Are the pass rates fair?
5. Do you/have you receive(d) any feedback in regards to the exams? If no, are you expecting any? If yes, when did you receive it?
6. Have you been given an opportunity to provide feedback on your examiners
7. Is the exam content appropriate for your level of experience?
8. In what ways (if any) do you think the exam could be improved?

Examiner Questions

- 1a) What do you think is the purpose of the exams?
- 1b) Do you think the exams meet that purpose? If yes, how? If no, why not?
2. Who are the intended candidates?
3. Is the exam mapped to a specific curriculum? If yes, which?
4. What is the evidence base behind the chosen components of the exam?
5. Are there any alternate assessment methods that could be used instead of the current system?
6. Is the current exam system reliable, valid, feasible and cost effective?
7. What is the rationale and methodology for setting a performance standard or passing score?

8. Is there any interaction between performance in different parts of the assessment? Is any of it weighted?
9. How is the security of test material maintained?
10. What feedback is given to trainees?
11. How were you recruited and inducted as an examiner?
12. Is there any on-going development or training for examiners?
13. Are there any processes or standards in place for marking papers?
14. How might the exam/assessments be improved?