

**The implementation of workplace based assessment in the
context of clinical radiology training**

William Hugh Ramsden

Submitted in accordance with the requirements for the degree of
Doctor of Education

The University of Leeds

School of Education

May 2014

The candidate confirms that the work submitted is his own and that appropriate credit has been given where reference has been made to the work of others.

This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement

The right of William Hugh Ramsden to be identified as Author of this work has been asserted by him in accordance with the Copyright, Designs and Patents Act 1988.

© 2014 The University of Leeds and William Hugh Ramsden

Abstract

Workplace based assessment (WPBA) was introduced into clinical radiology training in 2010. The purpose of the study was to investigate its implementation by addressing research questions concerning the assessments' day-to-day realisation, the participants' influence upon their enactment, and the utility of the process. The prior history of the introduction of WPBA into postgraduate medical training and the Royal College of Radiologists' guidance regarding the assessments' usage served as reference points for the study findings.

Semi-structured interviews were undertaken with 20 radiologists (12 trainers, 8 trainees) in order to discover their opinions of the implementation and usage of WPBA in radiology. The interview data were subjected to thematic analysis in order to identify issues of importance to interviewees, address the research questions and suggest any means of improving the assessments.

The analysis showed that most interviewees thought that WPBA had a formative purpose, as per central guidance, although some assessors felt it could be used summatively in certain circumstances. The day to day realisation of the assessments was subject to some variation, and although this might be seen as inevitable due to differing circumstances, there was evidence of both trainees and trainers manipulating the process to suit their own purposes. There was also evidence of some trainers frankly adapting the assessments' usage depending upon trainees' seniority or time in an attachment.

Interviewees described various weaknesses of the process, including the peremptory nature of some assessments, failure to identify underperforming trainees and poor assessor preparedness when WPBA was introduced. Reference to published literature from other postgraduate medical specialties showed that many of these problems were generic in nature, rather than confined to radiology. Suggested means by which WPBA could be improved included joint trainee and trainer leadership of the process, better training of assessors and refinement of the assessments themselves.

Contents

1. Background: Introduction of Workplace Based Assessment	1
a. Aims of Assessment in Postgraduate Medical Training	
b. Previous assessment methods and pressures for change	
c. Strengths and weaknesses of previous means of assessment	
d. Stimuli for change and aims of new assessment system	
e. Introduction of WPBA into Radiology and its planned purpose	
f. Summary of issues and outline of study chapters	
2. Literature Review and Research Questions	24
a. Literature Review	
b. Research Questions	
3. Methodology	43
a. The methodology of the study	
b. Rationale for the study sample	
c. Ethical issues	
d. Obtaining ethical approval	
e. Trialling of the interview schedule	

f. The method in practice	
g. Analysis of the data	
4: Introduction to Data Analysis	72
a. Selection of data analysis method used.	
b. Choice of themes.	
5: Theme 1: The purpose of workplace based assessment	79
a. Interviewees' views of the overall purpose of WPBA	
b. Summary	
6: Theme 2: How WPBA is realised in day to day practice	92
a. How the trainees approached WPBA	
b. Trainers' and trainees' actions as they prepared for an assessment	
c. Summary (a-b)	
d. How assessments were enacted	
e. How feedback was delivered	
7: Theme 3: How users influence the realisation of WPBA	134
a. Issues of adapting the assessments for summative use	
b. Recording WPBA	

c. Assessment of trainees without their knowledge	
d. Summary	
8: Theme 4: WPBA’s utility, weaknesses and suggested improvements	151
a. The utility of WPBA in assessing complex work	
b. Comparison with end of attachment assessment	
c. Summary of the utility of WPBA	
d. Weaknesses of WPBA	
e. Suggested improvements	
f. Summary of weaknesses and suggested improvements	
9. Discussion	176
a. The purpose of WPBA	
b. How WPBA is adapted for day to day use	
c. Power and control in WPBA	
d. Gaming and subversion in WPBA	
e. Summary	
10. Strengths and limitations of the study method	202
11. Conclusions	206

- a. Reflecting upon the research questions
- b. Implications of the study for future practice and research

12. References **219**

13. Appendices **229**

- a. Tools used in WPBA in Radiology
- b. Semi-structured interview schedule
- c. Participant Information Sheet
- d. Consent Form
- e. Numbers of WPBAs undertaken by respondents
- f. RCR Assessments (extract from curriculum)
- g. mini-IPX scoring form
- h. mini-IPX Guidance for Assessors
- i. Rad-DOPS scoring form
- j. Rad-DOPS Guidance for Assessors
- k. Data Coding System

Abbreviations appearing in the Thesis

1. ARCP: Annual Review of Competence Progression
2. BERA: British Educational Research Association
3. ERIC: Education Resources Information Center
4. GMC: General Medical Council
5. Mini-IPX: Mini-Imaging Interpretation Exercise
6. MSF: Multi-Source Feedback
7. NCBI: National Center for Biotechnology Information
8. NHS: National Health Service
9. NRES: National Research Ethics Service
10. OSCE: Objective Structured Clinical Examination
11. PMETB: Postgraduate Medical Education and Training Board
12. Rad-DOPS: Radiology Direct Observation of Procedural Skills
13. RCR: Royal College of Radiologists
14. WPBA: Workplace Based Assessment

1. Background: Introduction of Workplace Based Assessment

Assessing the competence and performance of those undertaking professional work has been a significant issue in many fields since the concept of professionalism was first acknowledged. This is particularly true of medicine, where practitioners' competence is critical to the health and safety of their patients.

In the pursuit of greater validity in the assessment of performance in everyday practice, a technique which has been increasingly applied is assessment of practitioners in the context of their workplace. This thesis concerns the introduction of workplace based assessment (WPBA) into the field of clinical radiology. It aims to explore the key themes of purpose, realisation, users' influence upon, and ultimately fitness for purpose of these assessments from the points of view of both trainers and trainees by means of an interview based study.

This opening chapter will introduce the reader to the underlying aims of assessment in postgraduate medical training, before describing some of the methods which have been used previously, including discussion of their strengths and weaknesses. The drivers for introducing new forms of assessment such as WPBA are outlined, prior to a description of its introduction into radiology. Finally, the chapter provides a rationale for the study and maps out its broad aims.

a. Assessment in Postgraduate Medical Training

Assessment in postgraduate medical training is long established, with a variety of means in long term use to try to ensure that doctors are competent and patients can have confidence in the practitioner who is treating them. Prior to discussing issues pertinent to assessment in medicine, it is important to note that there is a significant difference between formative and summative assessment, and as these terms arise throughout the thesis, they are defined at its start.

Both techniques are utilised in postgraduate medical training, and formative assessment may be regarded as assessment for learning, occurring throughout a course of instruction, with feedback on the learner's performance forming an integral part of the process. By contrast, summative assessment occurs at the end of a period of training, as assessment of learning, and enables a learner's performance to be graded or to have some other value judgement placed upon it (Rosenblatt et al, 2012).

Traditionally, formal assessment in post graduate medical education has been summative in nature (Burkill, 2008) and often based upon professional examinations undertaken away from the workplace. Trainees may only sit these examinations at limited times, and their high stakes nature is evident by success being a prerequisite to both progress through training and qualification as a trained practitioner. The examinations are conducted by the various medical colleges responsible, with the General Medical Council (GMC), for setting the curricula for postgraduate medical

training, in this study's case, the Royal College of Radiologists (RCR). The examinations and assessments mandated by the RCR are outlined in appendix f.

Allied to these professional examinations are the results of more locally based tests of competence, and less formal assessments such as comments passed by trainers upon trainees following an attachment or placement. If these are all satisfactorily completed, they lead to certification that a doctor has finished their training and reached the standard required for safe independent practice.

Maintenance of professional standards is cited within the wider aims of assessment discussed by Rowntree (1987) which also include selection, motivation of students, feedback to both students and teachers, and preparation for real life. However assessment relating to postgraduate medical trainees has additional purposes to that stated above, including helping their learning and development and the provision of evidence on which to base decisions regarding their progress through training, such as identifying those who are competent to pass onto the next stage (General Medical Council, 2010).

A previous working paper published by The Postgraduate Medical Education and Training Board (PMETB) (Southgate and Grant, 2004) was more specific regarding the purposes of assessment in post graduate medical training, adding informing career selection and driving learning to the attributes listed above. This working paper also listed further specific aims of assessment in postgraduate medical education, because of the high stakes involved in medical practice and the adoption of curricula based upon

trainees' demonstration of competencies in both their initial and specialist training (see below).

Aside from decisions regarding career progression, high stakes reasons for assessment included identification of trainees who should change specialty or even leave medicine, provision of evidence that training is complete, assuring patient safety, and providing evidence a practitioner may use towards revalidation. These aims are perhaps the most specific to postgraduate medical education, and appear to be largely driven by issues of patient safety and regulation of doctors' practice.

Because of the potentially important consequences of assessment in postgraduate medical education, it is important to have methods which are reliable and valid, particularly where issues such as certification to practice independently and patient safety are concerned. Reliability concerns how consistent and reproducible an assessment is when used repeatedly, whilst validity indicates whether or not it measures what it purports to. The assessment by means of a trial (as opposed to users' opinion) of the attributes of the WPBA tools proposed for use in radiology is included in section e.

Aside from the purposes stated above, a further use of assessment in postgraduate medical education is to support the use of new competency-based curricula by measuring trainees' progress in achieving relevant competencies. In this context, competence is defined as;

'a trainee's ability to perform a particular activity to the required standard (such as that required for patient safety), whilst being observed in the workplace' (General Medical Council, 2010).

Despite being titled a definition of competence (by the GMC) this statement actually describes the evaluation of a trainee's performance in the workplace, whilst competence is usually defined as the ability to undertake an action in an artificial test situation.

Traditional written exam-based forms of assessment are poorly suited to judging competence, as opposed to knowledge, and are even less well suited to evaluating trainees' actual performance in the workplace.

Although knowledge and elements of competence may be addressed by previous means of assessment, other attributes such as performance and action are not. Evaluation of the latter is further addressed in section e, and this shift in emphasis in what assessments are supposed to address has opened the way for new forms of assessment with higher claims to authenticity to be introduced into postgraduate medical education.

Because of this, WPBA has in recent years had an increasing role in medical training, with regard to judging competency and actual performance in the clinical environment. The guide for WPBA implementation issued by the GMC in 2010 identifies one of its strengths as; '*maps achievement in a competency framework*', (General Medical Council, 2010) and adherence to such a framework is now embedded in all specialty training curricula, including radiology. Within the curricula competencies are assigned to knowledge, skills and attitudes, and appropriate means of assessment, such as methods of WPBA (with expected outcomes), are mapped to these.

The GMC, which incorporated the functions of PMETB in 2010 is the regulator of medical practice in the United Kingdom, and the ultimate arbiter of how medical training is organised and delivered. Their authority with regard to postgraduate medical

training content, standards and assessment is devolved to the medical Royal Colleges, although these institutions still have to gain The GMC's approval for their curricula and means of assessment.

Though these organisations are the arbiters of central policy, it is individual trainers and trainees, who are members and fellows of the college, who have to enact the delivery of training and assessment in the workplace. Aside from their college's accountability to the GMC, there is also individual accountability, as the GMC is the registration body for all doctors. Those trainers with a mandate to organise or deliver training will have a large influence on how it is organised locally within hospitals and departments, and this feeds into one of the research questions outlined later.

There are other stakeholders in the assessment process, including patients and employers. Though they have definite roles in the provision of training, it is their interests which are most importantly served by the provision of appropriate training and assessment, for as was stated in the first paragraph, patients must have confidence in the practitioner who is treating them, and employers must be confident that doctors they employ are properly trained.

Following the GMC's instructions regarding the need for trainees to demonstrate specific skills to a required standard, the RCR issued a competency-based curriculum in 2010, with specific competencies having specific means of assessment mapped to them. Many of these assessments are undertaken in the workplace, and thus seek to address actual performance by the use of WPBA.

The assessments were introduced into radiology shortly afterwards (within the same year), with guidelines for their use provided by the RCR (discussed in detail in a later section). The RCR's curriculum and the guidelines (Royal College of Radiologists, 2010) were designed to inform a standardised process, instituted locally, by trainers and trainees throughout the United Kingdom.

The status of the guidelines (available to trainers on the RCR website) appears to be predominantly advisory, with terms such as '*please*' and '*should*' being used, rather than more prescriptive language. Interestingly, although the majority of the guidance in the College curriculum is couched in similar terms, it also states that; '*assessors must be trained in giving feedback and understand the role of the assessment.*' (their underlining) (Royal College of Radiologists, 2010).

How this centrally designed process was interpreted and subsequently realised locally when used by trainers and trainees is one of the main issues to be investigated by this study. The importance of this is that the RCR's policy advocates that WPBA is designed to be used in a certain way, but there are areas of the guidelines which are not explicit (or even ambiguous) which may be interpreted or influenced by users, leading to variation in how the process is realised. It might even be argued that such interpretive activity is an inevitable part of the implementation process.

b. Previous assessment methods and pressures for change

The professional examinations mentioned in the previous section are primarily measures of trainees' knowledge and clinical skills in an artificial setting, and as noted previously; do not assess their day to day performance. In a review of assessment in radiology, Augustine et al (2010) argue that although methods used in these examinations such as multiple choice questions are both reliable and feasible in the assessment of knowledge, they are not an acceptable means of addressing practical skills. Though not specifically mentioned in their paper, other attributes such as professional skills would also be poorly addressed by multiple choice questions (or any other written test) in comparison to WPBA.

Assessment of the latter has been primarily based upon the impressions of a trainee's supervisors, often resulting in a general comment at the end of an attachment or training period, rather than observation and assessment of specific activities. Additionally, ill defined attributes such as trainees' perceived attitudes and professional relationships might also be included within a general comment. Both performance in the workplace and the attainment of professional skills are now regarded as areas needing to be assessed specifically. In arguing for the latter, Francis (2008) suggests;

'Making professionalism explicit in permanent written record, which can be used to provide tangible and appropriately framed information for revalidation'.

Aside from a perceived need to better assess attributes such as trainees' performance and professional skills, there has also been an alteration of the training environment in which the assessment systems are embedded. Previously, training at this level was primarily based upon a system which might be termed apprenticeship, in which a trainer would spend sufficient time with a trainee in order to be able to make an informed assessment of their work.

However, the close contact between individual trainers and their trainees has been reduced due to the influence of changing work patterns such as shift working and a reduction in time trainees spend in the workplace. The reduction in trainer/trainee contact militates against one trainer being able to comment upon an individual trainee's work, and instead, trainees are assessed by multiple assessors with the aim of obtaining;

'snapshots of the trainee in the workplace are taken from multiple perspectives to build up an accurate picture of the trainee's performance at work'.
(Augustine et al, 2010).

Although reduction in prolonged contact between individuals may have the advantage of ensuring trainees are assessed by numerous trainers, rather than their assessment being dependant on the view of one assessor, current methods of assessment could be seen to be under pressure from other factors aside from a changed training environment which reduces time individual trainers spend with their trainees.

Additional areas of concern regarding established assessments include the attributes they are able (or not) to address and their difficulty in assessing performance in the workplace, both of which are discussed in the section which follows.

c. Strengths and weaknesses of previous means of assessment

One of the perceived strengths of formal examinations is a supposedly standardised means of assessment, in which all trainees aim to reach a particular level in order to pass. The high reliability of multiple choice questions (Augustine et al, 2010) as used in the RCR's examinations was noted in the previous section, and such assessments are subject to rigorous testing and peer review, in an effort to standardise the process and remove questions or other elements perceived as irrelevant or unfair.

Augustine et al (2010) define a reliable examination as one which will; *'identify competent radiologists consistently'* and it is also important that professional examinations demonstrate high validity by measuring those attributes critical to competent radiological practice. Prior to being used, exam questions are blueprinted to ensure that the curriculum is fairly represented overall, and to try to ensure that formal examinations sample all curriculum content and outcomes, (and are free of bias towards or against any subject area), but there may still be inconsistencies within the process.

Even written tests of background knowledge may favour certain candidates (or groups of candidates) by virtue of their content, style, and whether or not questions have been used before. An example of this might be a question on a specialist area of practice undertaken at some candidates' hospitals, but not others'. It must also be accepted that such tests can still only sample parts of a trainee's knowledge base without becoming unwieldy and impractical.

Written tests are not the sole method of assessment used in high stakes professional examinations, and many of these still contain an oral component. The RCR uses a viva as part of its final examination (but not in prior components) when a judgement is made regarding whether to award a trainee Fellowship of the RCR, which represents the basic postgraduate qualification in radiology which will be needed for future consultant practice in the NHS.

In this component of an examination, candidates would all be expected to view similar material of comparable complexity (as judged by groups of examiners) and be questioned about it in similar ways, within a broadly standardised framework. This has the aim of reducing subjectivity in the process, particularly when trainees are assessed by examiners they do not know, hopefully removing the influence of preconceived views.

Despite attempts to train examiners and weed out material considered unfair, it is unlikely that all sources of subjectivity can be removed from a face to face encounter, and it is probably unrealistic to try and eliminate them completely. Despite this, every effort should be made to treat candidates fairly, and individual assessors' subjectivity may be reduced by ensuring candidates are assessed by multiple examiners. However, it is very difficult to remove the influence of a personal encounter altogether, and this has been acknowledged by the American College of Radiology's decision to abandon the oral component of its examination (Wang, 2012).

In addition, the results of examinations are difficult to extrapolate to performance in the workplace, as highlighted by Swanson et al (1995), who made the point that

achievement (even in performance-based assessment) in one context is a poor predictor of performance in another. Even within the same context, it is impossible to know if a trainee's performance of one radiological skill (e.g. reporting a chest radiograph) reflects their skill at another (e.g. reporting an abdominal radiograph).

The outcomes of external examinations depend upon a trainee's efforts on the day, in a pressurised and artificial environment. An additional weakness of this system, depending on the actual assessment used, is the very limited feedback available to candidates (only in some circumstances, such as persistently failing an examination) upon their performance, with the loss of an opportunity to help them improve.

Previous informal means of internal assessment would ideally be based upon a large amount of contact between a trainer and their trainee, allowing the assessment to be based upon prolonged observation of clinical work with frequent feedback. This would allow facets of professionalism and day to day practice to be assessed in detail, complementing an examination system which would have difficulty in examining these areas.

Such informal assessment allowed assessors to pass informed judgement upon their trainees, and also allowed the assessor to smooth over the effects of particularly good or bad days, whilst acknowledging that a trainee's practice must always be safe. This means of developing an overall impression served to reduce the pressure on trainees to perform on the day, as would be required when sitting a professional examination.

However, this form of assessment can be criticised for lacking standardisation, with trainees unaware of the level of attainment required to pass. Trainees' assessments were

likely to be influenced by their trainer's personal preferences and prejudices, with the characteristics of adequate performance being an implicit rather than an explicit entity and other influences such as whether or not a trainer liked their trainee possibly also playing a part in deciding the outcome. As a result, a trainee's progress through training might be impeded by the opinion of an individual, potentially without the influence of other trainers' views.

This overall impression was also unlikely to be supported by verifiable evidence or documentation, making it very difficult for trainees to know where improvement was required, unless they received specific feedback. Such problems in the assessment of clinical practice were addressed by Moorthy et al (2003) who claimed that surgical trainers' views regarding their own trainees' technical competence were unlikely to be '*objective*' and would therefore be unreliable as a result. This may be true, but raises the question of whether any examiner can ever be objective, and does ensuring that a trainee is assessed by multiple assessors help minimise such effects?

In an overview of traditional assessment, Poikela (2004) argued that it was limited by being mainly concerned with the memorising of facts and the performance of tasks, and these are related to cognitive and operational processes. By contrast, social processes such as how learners act as a leader or member of a group, and reflective processes which might manifest as how they deal with problem situations and learn were less well addressed.

This view of traditional assessment could be applied to postgraduate medical examinations, reflecting their bias towards testing knowledge and problem solving,

rather than behaviours such as leadership and adapting to changing circumstances. Although they may be well suited to testing factual recall, such examinations do not address social and reflective domains as well, hence the need to develop alternative forms of assessment to complement them and widen the range of what is assessed. This may be regarded as contributing to the desire to improve assessment in postgraduate medical education, and the stimuli for this process are discussed in the next section.

d. Stimuli for change and aims of new assessment system

Some of the stimuli for augmenting the system of assessment in postgraduate medical education with WPBA stem from the perceived weaknesses of traditional methods outlined in the previous section.

A response to such concerns was contained in a working paper published by PMETB (Southgate and Grant 2004), which acknowledged the continuing importance of national examinations, but also indicated that there was a need for the assessment of doctors' performance in the workplace. This response also heralded the increasing use of competency based curricula in postgraduate medical training and the link between such curricula and the use of WPBA was made explicit in a subsequent report from the same organisation published in 2008.

The earlier working paper sought to institute an integrated assessment system to cover a junior doctor's entire postgraduate training, stating that, '*assessment will be based on curricula for postgraduate training*', and as such curricula specify particular skills trainees must be able to perform in the workplace, the argument was made for an

assessment system able to measure such performance. The importance of demonstrating specific competencies was one of the stimuli for the development of WPBA in medicine.

Although trainees might be asked to demonstrate evidence of specific competencies, there is an important difference between competence and performance. These two attributes form the top two levels of Miller's pyramid for assessing clinical competence (Miller, 1990), in which the author aimed to categorise means of assessment in terms of how well they could evaluate routine function. The lowest of four levels of the pyramid is assessment of knowledge, whilst the highest are described as '*shows how*' and '*does*'. The former measures competence in an artificial environment whilst the latter addresses authentic day to day performance and one of the main aims of introducing WPBA is to evaluate such routine function in the workplace.

Changes in the training environment (as alluded to in section b) have also provided stimuli for the introduction of WPBA. Trainees' hours are now limited by the European Working Time Directive, reducing the time for learning and that which they may spend with an individual trainer. This has been exacerbated by measures such as the introduction of full or partial shifts for trainees covering on call work, leading to a further reduction in trainer – trainee contact.

Southgate and Grant's discussion paper (2004) states that WPBA must be based upon evidence, '*collected and documented systematically*', and '*judged against pre-determined published criteria*'. This addresses (and aims to improve upon) the previously implicit standards required of trainees, by making them explicit and uniform

across training schemes, thereby providing documented evidence upon which trainers may base their comments, rather than an overall comment without such basis.

Thus the stimulus to introduce WPBA is seen to be multifactorial, including the need to assess trainees' progress through competency based curricula, increasing emphasis upon the assessment of performance in the workplace, changes in the work environment due to the reduction in trainees' hours and the perceived weaknesses of traditional forms of assessment, as discussed in the previous section. The process of introducing WPBA into radiology and the basis of its formative ethos are discussed in the next section.

e. Introduction of WPBA into Radiology and its planned purpose

The reasons for the introduction of WPBA into radiology mirror the general reasons for its adoption into postgraduate medical education, as outlined in the final paragraph of the previous section. In terms of radiology itself, Harding and McCoubrie (2009) highlighted that pre WPBA; *'many important aspects of radiological practice are not currently assessed in any meaningful fashion'*, citing an example of practical work undertaken in the workplace.

Despite assessments being mapped to specific competencies, suggesting that they might have some summative function, the National Pilot of WPBA in Radiology (Denison, 2010) emphasised that episodes of WPBA should be used formatively, but that a summative judgement might be informed by trends from multiple encounters. The tension between a formative and summative role for WPBA in radiology is one of the

key issues this study will address, and it may be argued that the trial documentation contributed to its emergence.

The national pilot ran from 2008-2009 and subjected the assessment tools proposed for use (appendix a), aside from multi-source feedback, to tests of educational impact, reliability, validity and feasibility. The data comprised 460 separate assessments (274 image interpretation and 186 procedural work) and associated comments from the users involved, and were subject to both statistical (e.g. reliability calculations) and qualitative analysis. The image interpretation assessments were undertaken by 113 trainees, whilst 95 undertook procedural work, though many individuals will be represented in both groups. In both cases the trainees involved were predominantly junior (years 1 and 2), and the majority only undertook one or two assessments in each category.

The image interpretation tool (mini-IPX) satisfied validity and reliability tests sufficient for it to be recommended for use in formative assessment, whilst the procedural tool (Rad-DOPS) performed well on some measures of validity but the pilot data were not suitable for reliability calculations, as trainee ability was found to have no impact upon the scores, when compared with assessor effects and variation between cases.

It was felt that the unsuitability of the data relating to Rad-DOPS for reliability calculations may be related to the small sample size and the fact that many trainees only undertook one assessment. It was considered that a larger sample might produce different results. This meant that the tool was not interpreted as unreliable, but continuing analysis of its use was recommended.

The formative ethos of WPBA was subsequently confirmed in the curriculum published by the RCR in 2010, which stated;

'Each WPBA should also be considered developmental and an opportunity for learning and feedback'. (Royal College of Radiologists, 2010)

No overt mention of summative usage was made in the curriculum, but later the document states;

'Formative workplace based assessments will enable overall competency and performance to be judged'. (Royal College of Radiologists, 2010)

This statement addresses the need for WPBA to address a competency based curriculum, and as a result blurs the boundaries between formative and summative usage, and although some overlap is perhaps not surprising, this forms a key issue of the thesis, as stated above. In addition the guidance issued on the conduct of individual assessments mandates completion of a form to document the episode in which the trainee's performance is rated against the trainer's expectations for their stage of training.

The ratings of trainee performance are entered against a checklist, followed by a rating of overall competence (ranging from 'requires additional support and supervision' to 'able to practice independently') and finally there is space for freehand comment from both the assessor and the trainee. All episodes must be followed by the delivery of feedback. The forms used to rate trainees' performance in the assessments (derived from

those used in the pilot) are included as appendices (Royal College of Radiologists, 2010).

The documentation used may also add uncertainty to the assessments' perceived purpose, by asking the trainer to score the trainee's performance. The ambiguity of not assigning either a definite formative or summative purpose to WPBA also pervades some authors' views on the subject in general, as evidenced by Norcini's (2007) view that; *'the methods bring together summative and formative assessment'*, suggesting that there can be a connection between the two. In the same review, he discusses numbers of cases required for assessment, suggesting that a collection of episodes, rather than an individual encounter, is needed to make a summative judgement, which represents a similar view to that expressed by the RCR trial.

The nature of the WPBA tools used in radiology are listed in appendix a. As will be seen from the list, all the methods aside from the audit assessment are based upon direct observation of trainees' performance whilst undertaking clinical work or teaching others. It is the usage of the two methods used to assess clinical episodes (mini-IPX and Rad-DOPS) which is examined in this study, and the RCR's guidance for the day to day conduct of these assessments is reviewed in detail at the start of chapter 6.

f. Summary of issues and outline of study chapters

The key issues surrounding the introduction of WPBA discussed in this chapter include

- Its use in the assessment of trainees' workplace performance

- Its formative or summative purpose
- Its day to day realisation
- The potential for adaptation by end users

Although these represent topics important to WPBA in general, they are also pertinent to the assessments' use in radiology, which also has more specific issues including how users in the specialty have realised WPBA locally and its use in assessing complex work. All of these issues are addressed in the study.

Throughout postgraduate medical education, issuing of curricula containing specific competencies which trainees must achieve was an important driver for the introduction of WPBA, but the difference between the assessment of competence and performance is important, and sometimes not fully addressed. The document 'Standards for curricula and assessment systems' (Postgraduate Medical Education and Training Board, 2008) relates the former to national examinations and the latter to WPBA, following this by stating;

'Competence (can do) is necessary but not sufficient for performance (does), and as trainees' experience increases so performance-based assessment in the workplace becomes more important.'

With regard to the assessment of performance, the quotation suggests that the need for this becomes greater as trainees' experience increases, yet this factor is not mentioned elsewhere, and, ironically, the system of WPBA was first used to assess the most junior trainees of all. Grant et al (2007) argued that the assessments were more appropriate at

lower stages of training, and this would suggest that the rationale for their introduction is no less compelling for junior trainees.

The documents which introduced WPBA do not state whether it was primarily designed to be used formatively or summatively, although both of the cited papers published by PMETB (2004 and 2008) are explicit regarding the need for feedback afterwards. This would imply that a formative element was always intended to be included in WPBA, in addition to the assessment of performance (and the reasons for doing this) referred to earlier.

Although usage in both formative assessment and the assessment of professional skills (as referred to previously) may have been part of the general rationale for the introduction of WPBA, it is difficult to find explicit evidence for this in advance of its introduction. A danger of a 'retrospective' view of prior reasoning is that one may be influenced by how the assessments have come to be used in practice, and presume that their current usage was pre planned, when it actually represents a subsequent development rather than the rationale for their introduction.

If subsequent developments are ignored, the main external reason for the introduction of WPBA was the introduction, and now widespread use, of competency based curricula, with the requirement that assessment systems are appropriate to demonstrate (and document) that doctors have acquired these skills. The latter implies assessment of actual performance, consequently situating at least part of the process within the workplace. This was driven by PMETB, now subsumed into the GMC.

With regard to the assessments' usage, it would seem that organisations such as the RCR have primarily adopted WPBA for formative use, allowing for the issues regarding its purpose raised in the previous section. The intended purpose of the assessments (formative and/or summative) derived from statements in the college's trial of WPBA and the subsequent curriculum raise important issues which are further pursued in the study. As will be seen in subsequent chapters, this is perceived as a real, rather than a potential problem in other specialties and it is important to see if this situation pertains in radiology. It is also important to realise that users may consider that WPBA has other purposes as well.

Users' perceptions of WPBA's purpose and their engagement with the process are essential factors in deciding how the assessments are realised day to day. An assessment may be regarded as an encounter constructed by the trainer and trainee, and their views, how they interact, and the authority relationship between them are all vital ingredients in determining how the process is enacted. Exploration of the issues surrounding the day to day realisation of WPBA forms a large part of the study.

Closely related to end users' realisation of the assessments is how they may influence the enactment of WPBA, leading to adaptation of the process when compared to central guidelines for its usage. Users may choose to adapt or ignore central guidelines, if their personal view of the assessment process is different and they see an opportunity to adapt it to their own purposes. This issue is also explored in the study, with particular reference to some users' manipulation of WPBA.

In addition, how the process is realised and potentially adapted are likely to have a major influence on users' view of its utility, particularly whether or not WPBA can contribute to the assessment of complex work. An evaluation of WPBA's utility could be undertaken for its own sake, but when it concerns an evolving process which has only recently been introduced into radiology, it would seem natural to try and see if there are ways in which users feel it may be improved. This issue and those discussed in the paragraphs above inform this study's research questions, and these are discussed further in the next chapter.

This 2nd chapter (literature review and research questions) serves to review literature pertinent to WPBA from postgraduate medical education and other contexts, prior to outlining the research questions. This is followed by the 3rd chapter describing the methodology of the study, including discussion of the most appropriate means of addressing the research questions.

An introduction to the data analysis comprises the 4th chapter, including discussion of various methods of analysing the data, and justification of the selected method, prior to four chapters (5-8) of data analysis. These are followed by the discussion (chapter 9) and a short chapter reflecting upon the methodology used in the study (chapter 10). The concluding chapter (11) follows, including reflections upon the research questions, and the implications of the study's findings for future practice and research. A list of references and appendices a-k complete the thesis.

2. Literature Review and Research Questions

a. Literature Review

Much of the background literature concerning WPBA relates to its use in undergraduate and postgraduate medical education and although review of this material forms the majority of this section, examples of the assessments' usage and policy adaptation in other domains are also discussed. This review focuses on issues raised in the literature relevant to the study including the assessments' role in evaluating performance, their purpose, how users engage with them, their perceived usefulness (particularly with respect to assessing complex work), and how central policy, both in WPBA and in general, may be adapted by end users. These topics are used to inform the study's research questions, which are introduced in the second section.

Material for the literature review was predominantly sourced from The National Center for Biotechnology Information (NCBI) PubMed and Education Resources Information Center (ERIC) online databases. The basic search term used was workplace based assessment, and I had to specify the inclusion of all three words to ensure the results remained manageable. Specific terms such as radiology, adaptation and policy were added when I wished to address particular areas. I also obtained material by more informal means, including cited papers' lists of references, supervisors' and colleagues' recommendations and more general online resources such as Google Scholar.

Despite using specific search terms, the databases yielded a plethora of material and inevitably I had to be selective and include the literature most relevant to the study. This

comprised both empirical studies and other material concerned with the subject, including policy documents, discussion papers, opinion pieces and review articles.

The assessment of competency and performance in the workplace context is very important. The need to evaluate professional performance in an appropriate context is supported by Eraut's (2004) view that; *'judgements of competence are still very situation specific'* and despite the use of a different term (competence) it is clear that he views the context of a practitioner's actions as integral to how well they are performed.

The great importance of the context in which assessment occurs is also emphasised by Govaerts and Van der Vleuten (2013) who argue that performance in WPBA may not be a stable entity, and is highly dependent upon the environment which surrounds the assessment episode. Assessee's performance is felt to be dynamic, affected by environmental factors, and based upon social and cultural circumstances, as well as the assessment's context. Thus notions of performance as a fixed entity measurable between contexts may be flawed. The authors also challenge the assumption that competence is a fixed, permanent attribute and both it and performance are seen as variable and dependent upon circumstances which may be unique to particular situations.

Thus the difference between competency-based assessment and performance-based assessment described by Rethans et al (2002) as the differences between what doctors are able to do in high stakes controlled conditions and what they do in day to day practice may be true, but the variation may be dependent upon many other factors in addition to an assessment's context, as suggested above. Rethans et al's (2002) conclusions are based upon very variable correlation between doctors' results in external

examinations and their practice in the workplace, and it is the latter which is claimed to give WPBA its unique status in evaluating performance.

Outside medicine, the importance of context and having a realistic setting for the evaluation of performance is cited by Fox et al (1998) who evaluated the assessment of driving competence after brain impairment. They found that although an off-road assessment could evaluate drivers' proficiency in operating a car, an on-road test was needed to properly assess driving in traffic or to predict future safe practice. The latter is particularly pertinent in medical WPBA, as the link between such assessments and prediction of future performance is critical for patients' well being.

Based upon Fox et al's (1998) work, the overall purpose of WPBA may thus be defined as realistic evaluation of routine performance in a 'natural' environment (prior to any assignment of any formative or summative purpose), but factors other than the setting are also important, such as the number of observed episodes needed to make an accurate assessment. Norcini (2007) notes that judgements based upon a doctor's performance in a single encounter may not accurately predict performance of another task, and the suggestion that multiple episodes of WPBA may be required to accurately assess average performance should be considered in addition to the need for an appropriate context cited by Eraut (2004).

As a result of this Norcini (2007) advocates that assessments are undertaken over a period of time to determine routine performance (and thus increase their reliability), reducing the case specificity of an isolated episode, which might otherwise skew the overall result. Factors such these serve to highlight the issues that surround undertaking

WPBA, and what needs to be taken into account to try to ensure the satisfactory assessment of day to day practice.

The need for adequate sampling to provide reliable assessment data was emphasised by Wilkinson et al (2008), who reported the results of a trial of WPBA on behalf of the Royal College of Physicians, stating that at least twelve episodes were required to provide data of acceptable reliability. This was based upon the analysis of two hundred and thirty medical trainees' completed assessment forms (and an additional questionnaire) followed by the use of generalisability theory to model the scores' reliability in terms of different numbers of WPBA episodes and assessors.

Beyond the need for adequate sampling to provide reliable data, Grant et al (2007) found that when WPBA was utilised in the Foundation Programme to assess junior doctors, its interactive nature around real cases led to confusion as to whether these episodes represented assessments or educational events. If the assessment is used formatively, it could represent a combination of the two, but the uncertainty over its purpose affected the trainees' approach to these episodes, raising the possibility they would prepare specially, causing doubt that their normal day to day practice was being assessed.

Writing about WPBA in medicine, Norcini and Burch (2007) emphasise its formative purpose, and its strong association with the provision of feedback. They describe the process as an important means of changing learners' behaviour to achieve desired outcomes, and thus it is possible to link this to the more fundamental purpose of assessing routine function in the workplace.

Norcini and Burch's (2007) view is supported by Boud (2000), who argues that sustainable forms of assessment are needed to underpin lifelong learning, and that if this purpose is fundamental, the emphasis of WPBA should be on its formative use, rather than summative evaluation of performance. Brodie and Irving (2007) report upon the usage of such a strategy in the assessment of work-based learning and advocate that assessment tasks should form an integral part of the learning process.

Norcini and Burch's (2007) view of a well understood formative ethos of WPBA is at variance with others' opinion that its purpose has often been misinterpreted by users, leading to its usage as a summative evaluation of performance. This may have occurred in the Foundation Programme (the initial two years of postgraduate training undertaken by all doctors immediately post graduation), in the opinion of Kessel et al (2012) who suggested that the planned purpose of WPBA was often adapted, stating that;

'Unfortunately, the purpose of these assessments in the foundation programme was often misinterpreted, and they have been used mainly as a summative assessment of outcomes of learning rather than as a formative stimulus for further learning'.

A further argument against such assessments being graded or summative was advanced by Sadler (2010), who suggested that when assessment is used in this fashion, the results are incorporated as credits into the student's record, signifying a sense of closure of the assessment episode. This militates against the assessment being used to stimulate further learning and may encourage grade- seeking behavior when the reward sought by the student is a good result rather than useful feedback. By contrast, if there are no grading

stakes in an assessment, this frees up the educational environment, allowing both teachers and students to be imaginative and explorative in building future learning.

Surveys of users in higher training demonstrate further contrasting views of the assessments' purpose. A questionnaire survey of Psychiatry trainees' opinions of WPBA undertaken by Menon et al in 2009 found that 43% of respondents felt the assessments were introduced '*to improve training*' whilst 41% perceived their introduction as '*politically driven*'. In addition their survey found that most trainees felt that the assessments had no real benefits upon their training, clinical practice, supervision or confidence.

Although the results of the survey may predominately reflect the perceived impact of WPBA, rather than its purpose, it might be argued that the majority of the respondents felt that WPBA has very little purpose in training. This was not an isolated finding amongst one group of trainees in psychiatry, as a survey of other trainees in the same specialty (Babu et al, 2009) found that only 39% of those surveyed found WPBA useful in the development of their clinical skills.

These rather negative findings about experiences of the process might be seen as one extreme of how users perceive the utility of WPBA, and by contrast other authors portray the assessments as reliable and useful indicators of performance. The latter view was articulated in material cited earlier in this chapter by Wilkinson et al (2008) following a trial of WPBA on behalf of the Royal College of Physicians, based upon the analysis of both assessment score sheets and questionnaires. The trial findings described many positive comments regarding WPBA's formative value from analysis of the

questionnaires, and it is interesting to note the apparent divergence of views between those who are assessed and those who administer the system.

This divergence of trainees' and trainers' views in published material upon WPBA demonstrates their differing perspectives of the process, and their contrasting opinions are revisited throughout the data collection. Despite their differing perspectives, some interviewees advocated shared trainer and trainee leadership of WPBA, suggesting a desire for a degree of joint ownership of the process. Although it is unlikely that assessors and assessees could ever be regarded as equally authoritative owners of an assessment system, joint engagement might be seen as a means of the two groups working together to optimise the process, and possibly reduce potential manipulation or abuse of WPBA by either party. The subject is further discussed in chapters eight and nine.

When considering the different views of trainers and trainees regarding the utility and overall purpose of WPBA, it is not surprising that the junior trainees reported upon by Grant et al in 2007 were confused about the assessments' purpose in the Foundation Programme, leading these authors to conclude that there is; *'obvious blurring of the boundaries between assessment and education'*.

In addition to trainees' possible confusion between assessment and education, some studies highlight the impact of alternative perspectives upon its purpose. Thus a report issued by The Academy of Medical Royal Colleges in 2009 mentions that the existence of numeric targets of assessments for trainees may mean that some may just view their

purpose as hurdles to jump rather than an important facet of training which may help guide personal development.

Although Wilkinson et al (2008) found that WPBA represented reliable and useful indicators of performance, it is important to try and understand if this could also be applied to radiology, and the paragraphs which follow discuss the ability of WPBA to address complex medical activity, which represents a feature of specialty training in all postgraduate medical disciplines, including radiology. The inclusion of this issue was stimulated by Grant et al's (2007) findings of significant concerns amongst users that WPBA was only suitable for assessment in the lower stages of training, where the evaluation of complicated clinical work was less important than the assessment of simpler, more isolated competencies.

The assessment of complex procedural work has already been addressed in specialties other than radiology, with varying findings. Moorthy et al (2003) investigated the assessment of technical skills in surgery by multiple means including checklists, global ratings scales, and dexterity analysis systems, concluding that such methods are; *'suitable for the objective formative feedback of technical skills during training'*.

However, a contrary opinion with regard to WPBA's ability to assess the overall performance of complex procedural work was advanced by Setna et al (2010), who noted that observed skills were often divided into smaller components during WPBA, and they queried whether the sum total of these small steps could really equate to overall competence. It might be possible to perform all steps in a procedure or interaction correctly, yet undertake them in an uncoordinated way or even the wrong order.

The same authors concede that although checklists might be reliable in the assessment of technical skills, more complex issues such as assessing learners' interactions with patients might be better assessed by using global rating scales. The implication that such behaviours may be more difficult to assess than technical skills is supported by Gaba et al (1998) who studied the assessment of both in the context of anaesthetists' performance during simulated crises. Although assessors achieved high levels of consistency in scoring technical skills, their views of how the assessees behaved whilst trying to manage a crisis were far more variable.

Discussion regarding the appropriateness of WPBA to evaluate complex activity is not confined to medicine, and some of the issues raised in the preceding paragraphs have been debated in other fields. With regard to whether the achievement of individual steps in a process can indicate overall competence, Hamman (2004) cites examples from aviation simulator training to argue that a complex scenario can be reduced to smaller 'event steps' and that achievement of these can indicate overall competence.

This view contrasts with a view expressed by Stanley et al (2002) regarding the assessment of music performance. Although some examiners felt that steps or criteria relating to them were helpful in assessment, allowing them to focus on important issues and helping them feedback to students, others felt that criteria-based assessment was too narrow a means to assess a student's performance, and that an overview was required.

The potential limitations of basing assessment upon multiple criteria rather than a holistic overview were also highlighted by Sadler (2009), who felt that such an approach was not compatible with the full judgement required by complex performance. He cited

problems such as whether criteria used were always truly representative of the most important attributes of a student's work and of examiners who might take a view of performance overall, then retrospectively fit in appropriate criterion-based marks to suit their holistic view of the whole, which might be greater than just the sum of its parts.

Although criterion based assessment might appear more 'objective', it may lead to a more mechanistic evaluation of a trainee's work, and in Sadler's (2009) view, it cannot address all the nuances of complex performance. This opinion was supported by Yorke (2009) who viewed the assessment of complex work as a judgement (as opposed to a measurement) which cannot be forced into a specified list of assessment criteria. The use of criteria thus might be seen as atomising the assessment of professional work with the potential detriment of not taking an overall view of performance.

In another example from outside medicine Gonczi (1994) cites the example of competency based assessment being utilised to determine the accreditation of solicitors in New South Wales. The assessment is partly based upon simulated practice in the workplace (interview with client, assessing mock legal files) and although based upon simulation, the aim of creating a realistic environment means there are some similarities to WPBA in postgraduate medical training.

The stated purpose is to help the public and fellow professionals identify solicitors with special areas of expertise, with no mention of helping develop such skills by the delivery of feedback, identifying the assessments as purely summative in nature. The use of such assessments to determine accreditation (rather than as part of training) suggests that in this domain, there is confidence in their ability to assess complex work.

Delandshere and Petrosky (1998) state that important aspects of performance vary according to the time and the context in which they are assessed, as suggested by Eraut (2004) and Govaerts and Van der Vleuten (2013) at the start of this section. The theme was further developed by Govaerts et al (2007) who argued that those who assess performances are not passive measurers of it, but instead more active participants whose behaviour is affected by both the assessment's context and their interaction with the trainee, suggesting that the two parties build the assessment episode together. As a result, the authors recommend that assessors' training should cover both assessing and the assessor's motivation during the process, as the latter will have a great influence on how the assessment is constructed between the assessor and trainee.

Variations in assessors' personal epistemologies may have a great influence upon assessments. Work from the department cited in the previous paragraph (Govaerts et al, 2013) concerning assessors' influence upon WPBA states that; *'raters make and justify judgments based on personal theories and performance constructs'*, and that raters often used personal schemas (allowing efficient information processing) when assessing performance. The authors found that such personal views often confounded attempts to train and standardise raters' assessment practice, but allowing groups of assessors to define performance levels themselves, within frames of reference, led to the development of shared values which could then be applied to WPBA.

Lack of success in targeting perceived rater subjectivity was also discussed by Gingerich et al (2011) who suggested that assessors tended to categorise assessees in an idiosyncratic fashion, but that overall, their impressions tended to consistently fall

within a few overall groupings rather than be completely disparate. Regarding such assessor variation as error is regarded as too simplistic by these authors and they suggest that if categorical judgements of competence or incompetence can be incorporated into an assessment system, this may be more reliable than trying to translate them into a mark or hierarchical grade.

These examples show that the issue of the utility of WPBA in assessing complex activity (technical skills and behaviours) remains contentious in many ways, including whether it is possible or not, and if it is, how it should be undertaken, how it should be measured or rated (checklist and/or overall impression), and the influence of the assessor upon the process.

The concept of assessors being active participants rather than passive measurers of performance during assessments noted by Govaerts et al (2007), suggests that trainers and trainees construct WPBA episodes together, and thus end users are likely to influence the conduct of the resultant assessment. There may even be frank adaptation of central policy due to both end users' actions and local circumstances.

Outside medicine, the importance of end users to policy implementation and the belief that some amendment of central policy is almost inevitable due to varying local circumstances is articulated by Darling-Hammond (1990) in a discussion of instituting teaching policy in schools, stating;

'local leadership and motivation for change are critical for policy success; that local ideas and circumstances always vary (therefore local agencies must adapt policies rather than adopting them).'

This view is supported by Younis and Davidson (1990) who described an overriding view of the tension between those who make policy and those who enact it (with respect to the realisation of policy in general), stating;

'The process may be seen from either a bottom-top or top-bottom perspective; policymakers will make decisions which will attempt to limit the power of other actors; actors will make decisions which will evade the power of decision-makers.'

Thus in many contexts, researchers have argued that it might be viewed as inevitable that central policy will be locally adapted, particularly if it is delivered as guidelines which cannot address every eventuality or local circumstance. Prior to any utilisation or conscious alteration of a policy by users, it is subject to; *'inferential processing, translation and interpretation at very many points'* (Dunsire, 1990), leading to the potential for conscious or unconscious amendment at multiple levels, and this author concurs with Darling-Hammond in suggesting that differences between intention and output are very likely.

The influence of guidelines upon policy implementation is addressed by McLaughlin (1987) who argues that if such guidance is weak, dominant groups or competing issues may shape what actually occurs, whilst Darling-Hammond (1990) emphasises the desirability of informing and supporting end users tasked with instituting new policy.

In a discussion of the relationship between policymakers and the end users who have to implement the policies, Chrispeels (1997) analysed the institution of new systems of student evaluation in schools in California over a ten year period, and found that in addition to top down dissemination of policy, policies were also influenced by feedback

from end users. Thus in addition to just being enactors (and possibly adaptors) of policy, the end users' role was enhanced by their potential influence upon future iterations of it. The author also noted that this was a dynamic and shifting process, as priorities and political leadership changed.

The importance of end users in instituting central policy in another context (mathematics education) is described by Cohen and Ball (1990) who on one hand describe teachers' practice as the problem the policy is trying to correct, whilst the on the other acknowledge that the teachers are the most important factor in improving things.

An example of adaptation of central policy with respect to assessment in radiology prior to the introduction of WPBA was published in *Clinical Radiology* (the College's own journal) by Long in 2001. The author (a local training programme director) described how the RCR's assessment form had been locally modified to try and improve it. This was undertaken by adding features such as space to record trainee appraisal and objective setting at induction, making it more subspecialty specific, and requesting trainees' feedback upon their attachments. These features were subsequently included in the RCR's own assessment forms.

This is an example of assessors adapting central policy for (in their view) the benefit of training, arguably validated by the subsequent incorporation of the changes by the RCR. Though deviation from central policy might be viewed in a negative light, it is not always the case, and it is end users who have the most intimate view of WPBA, which may work to their advantage when trying to optimise the process. An example of this

might include an assessor emphasising the positive aspects of a variable performance in order to try and boost the confidence of a struggling trainee.

This example of professional discretion being utilised to (hopefully) optimise WPBA, is characterised by Sadler (2009) as an example of a bestowed credit, where an assessor exercises their prerogative to encourage the assessee, without the latter engaging in deliberate behavior to earn such a reward. The study contains examples cited by interviewees of assessors' discretionary behavior having perceived positive and negative effects upon the process, and the subject is further developed in chapters seven and nine.

The converse of this is transactional credit where incentives and rewards known to both assessor and assessee prior to an assessment may be 'traded' during the episode, potentially encouraging particular behaviours on the part of both parties. As stated earlier, if an assessment is graded, such transactions may encourage grade-seeking behaviour on the part of the student, rather than seeking useful feedback upon their performance.

Overall, this section has reviewed whether the results of assessments are transferable between contexts within the same professional field, discussed the importance of assessment being undertaken in a natural environment, and considered the need for multiple episodes in order to evaluate day to day performance. It has used examples from outside medicine to look at the evaluation of complex activity in different contexts and how central policy may be influenced and adapted by end users.

It has considered the views of both the instigators and trainees regarding the assessment process, noting the dissatisfaction expressed by some users. Major issues revealed by

Grant et al's (2007) study of the use of WPBA in the Foundation Programme, such as confusion regarding the assessments' purpose, their suitability for assessing complex work and trainees' differing approaches have also been reviewed, and these are revisited later in the thesis and compared with the study's findings.

b. Research questions

The previous section examined the literature behind the current study and looked at some of the issues which underlie the introduction of WPBA. In this section these issues are utilised to inform the development of the study's research questions. It begins by discussing day to day enactment of WPBA, continues with regard to how end users may influence central policy, prior to stating the importance of investigating users' views of the assessments' utility and how they feel the process might be improved.

Aside from the uncertainty of some users regarding the assessments' purpose mentioned at the end of the previous section, it is important to remember that individual episodes of WPBA will be subject to many variables influencing how trainees and assessors realise an assessment, including the difficulty of the case, whether or not the assessor has assessed the trainee in the past and the seniority of the trainee.

These factors will help determine how the episode is constructed, how useful it is perceived to be (both to the trainee and others), and whether it is regarded as a low stakes assessment to aid personal development or a more summative evaluation of performance in the workplace, signifying an achievement if it is 'passed'. There are

also strong external influences upon users' participation in WPBA, particularly trainees, for whom the scheme is compulsory.

Thus trainees' and trainers' perceptions of how they participate in and enact WPBA are important, in terms of both how assessments are constructed and how they are utilised in radiology. The first research question seeks to utilise participants' accounts of the process to ask;

1. How has the policy of introducing Workplace Based Assessment into clinical radiology been realised in practice?

As exemplified in the previous section, there may be some tension between centralised policy and how a process is enacted locally, and a similar situation may pertain to the institution of WPBA in radiology. In this context, the policymaker is the RCR, whilst the local realisation of WPBA is the responsibility of practising radiologists, all of whom are members or fellows of The College.

It is important to note that some flexibility in policy may be essential, as the circumstances and resources of hospital departments and training schemes will inevitably differ. This has to be taken into account when examining how WPBA is enacted, as what may be possible in one place may not be possible in another, and as noted in the prior section, there have been instances in which assessment forms were previously locally adapted.

This adaptation was undertaken at training scheme level, and must be differentiated from more fundamental changes or amendment instituted by end users themselves,

perhaps defined as those which contradict or undermine College policy regarding WPBA, which may have drastic effects upon the assessment process, against the wishes of its creators.

Local changes to WPBA policy may be due to deliberate actions to enact assessments differently (either by adapting central guidance or exploiting gaps in it) and instituted by either organisations (e.g training schemes) or individual end users. Alternatively, or in addition, imperceptible amendments may occur as the policy is handed down (multiplied by the number of levels through which it passes). Thus, the resulting changes may be multifactorial, and act in partnership with each other to produce an assessment process different to that envisaged by policymakers.

Opinion regarding the local realisation of central policy in the previous section suggests that alteration of it by end users may be very likely, and the local enactment of WPBA is likely to be influenced in a similar fashion. Thus it is important to understand that when WPBA has been realised locally, whether central policy has been adapted, and if it has been, to what degree. Analysis of this issue informs the second research question, which asks;

2. How do the different participants engage with, and influence the enactment of, the Workplace Based Assessment process?

Although the second research question is distinct in terms of addressing participants' influence upon the assessments' enactment, there is some overlap with the first, as both concern the realisation of assessment in the workplace. A related issue to the realisation

of the assessments is their perceived utility, as whether or not users find the resultant process useful is crucial as to whether or not WPBA is retained and utilised in radiology.

This is an issue deserving of further enquiry, not just for its own sake, but because of concerns raised from Grant et al's (2007) evaluation of WPBA's utilisation in the Foundation Programme, which suggested that such assessments were only suitable for the lower stages of training, rather than the assessment of complex activity. They commented that despite this, the assessments were being implemented at all stages of training, making evaluation of their perceived utility with regard to more complex work (as found in radiology) essential.

It was thus felt important that the interviewees were asked directly about the utility of WPBA in the assessment of complex work, and this was linked to enquiry regarding the delivery and impact of feedback, as users' views regarding the success (or otherwise) of these processes would represent significant indicators of the scheme's suitability for radiology.

As WPBA had been newly introduced to radiology when the interviews were conducted, and it may be considered as an evolving process, users were also asked how they thought it might be improved, and this enquiry has been added to that regarding its utility, resulting in the third research question, which asks;

3. How do participants perceive the utility of the system of workplace based assessment which has been developed and how do they think it could be improved?

3. Methodology

a. The methodology of the study

The data needed to address the research questions outlined at the end of the last chapter would have to be sourced from the different groups who utilise WPBA and its results.

Although those who use the results might include various groups, such as the GMC, the RCR and hospital managers, information regarding the realisation of the process is most likely to be obtained from those directly involved, namely trainers and trainees.

Due to their differing positions within the process, both groups are likely to have distinct perspectives regarding their contributions to the realisation of the assessments, and because the questions largely relate to users' perceptions of, engagement with, and influence upon WPBA, the method of data collection should allow detailed exploration of participants' experiences and views.

In view of this, it was felt that the research questions would be best addressed by a method which would allow the pursuit of appropriate data and its subsequent interpretation without sacrificing subtlety or obscuring minority views. The data required to address the research questions would be derived from users' opinions and thus the methods which were considered were those in which their views would be sought directly.

There are various methods which might be used collect data to address the research questions, including interviews, questionnaires and focus groups, and the advantages

and disadvantages of each are discussed in turn. Addressing the research questions required information to be gathered in depth, including exploration of the reasons why study participants may hold particular views.

Interviewing

The most appropriate means of data collection appeared to be interviewing, and this was the method I utilised. It has the advantages of allowing areas of interest to be pursued in depth, the reasons behind interviewees' views explored, and the flexibility to be adapted to new issues arising as the study progresses. The method has the potential to yield rich data (DiCicco and Bloom, 2006) and if semi structured interviews are used, a degree of standardisation may accompany data collection, allowing respondents' views to be categorised and themes more easily derived from them.

A disadvantage of the method is that it is very 'user dependant'. This commences from the writing of the interview schedule and the need for the researcher to ensure it addresses the research questions in an open and non directive way. Similar issues pertain with regard to the schedule's utilisation by the interviewer, and how the data are interpreted. If the interviewer's technique is poor or there are previous issues between interviewer and interviewee, the resultant data may be poor, incomplete or compromised in some way, though this might be said of the utilisation of other methods as well, such as suboptimal facilitation of a focus group.

In the context of the current study, where trainees are to be interviewed by an individual with some managerial responsibility for them, there are also significant ethical issues

with regard to the authority relations between the two parties, and these are discussed in later sections. A further disadvantage of interviews (when compared to methods such as a questionnaire survey) is a relatively small sample, due to both potential respondents being unwilling to commit to a time consuming and intimate process, and practical constraints, such as the number of interviews a single researcher can conduct and analyse.

When these disadvantages are considered, it is clear that the use of interviews involves some compromises, such as the inclusion of data provided by trainees who might be wary of (or seeking to please) the interviewer, and a smaller sample size than might be possible with another method. Such issues must inevitably be considered when the data is analysed, and appropriate caution exercised (different methods of data analysis are discussed in chapter four).

Despite issues such as these, interviews appeared to provide the best means of obtaining the data which would address the research questions. The semi structured format adopted meant that interviewees were initially asked standard questions, but were allowed (and encouraged) to develop their answers more fully as individual topics were discussed, allowing significant issues to be pursued.

The study sought to recruit twenty subjects, and although another method might have gained more respondents, the data obtained (e.g. from questionnaires) may not have been as detailed, and twenty interview transcripts would be likely to contain much data. The development of the interview schedule is discussed later and that used in the study is included as appendix b.

The schedule was refined by undertaking three trial interviews, though an alternative method of doing this would have been to have used focus groups to define which issues were important. Potentially, the latter could have been a good means of clarifying and adding issues, and one might get an overview of their possible significance by seeing if they represented majority or minority views. However, once respondents had provided their opinions in a group situation, it would have been difficult to subsequently interview them when their views would have been coloured beforehand, and I felt I did not have a large enough pool of potential respondents to be able to find a new group of interviewees.

Thus the trial interviews (discussed in a later section) proved a successful means of refining the interview schedule and including new issues without using up large numbers of potential interviewees.

Interviews were used throughout the study. A refinement which I considered was returning the interview transcripts to interviewees after the episode to see if they wished to correct them. I decided against this as I felt prolonged consideration of the transcript might reduce interviewees' candour, they might not welcome the opportunity to have to read between thirty and forty pages of type, and some might not respond to the repeat invitation.

Although interviews were considered the most appropriate means of data collection overall, it is important to remember that a semi structured format only represents one way in which they may be used. Consideration was also given to structured and unstructured interview formats.

Unstructured interviews would have allowed more open pursuit of issues which were important to individual interviewees. However, the unstructured format would have made the data analysis (and subsequent derivation of themes) more difficult, and there may also have been issues concerning the reliability of the data collection, had there been great variability between interviews and no standardisation. There would also be the chance of an important area being left uncovered, were the interviewee not prompted to discuss it.

Semi structured interviews allowed both a degree of standardisation and some flexibility to pursue individuals' issues. Use of a structured interview format would have negated this flexibility, and it would have been difficult to pursue individuals' issues in depth. A structured format might have led to interviewees' responses being constrained by their having to fit into predetermined categories, with the loss of subtle data (Cohen et al, 2000), so consequently this was not pursued.

Questionnaires

As stated earlier, questionnaires were also considered as a means of collecting data, which could subsequently be subject to either qualitative or quantitative analysis. The use of questionnaires has the advantage of allowing wide coverage of a subject and the potential to reach a large number of possible respondents. There is also a lessening of interpersonal issues such as authority relationships which may affect other means of data collection such as interviewing, meaning that subjects may feel able to be more candid in their responses.

However the disadvantages include reduced potential to pursue areas of special interest which might arise, inflexibility in pursuing issues such as why respondents might hold particular views, and inability of the method to adapt to either individual respondents or the group, should early responses suggest unforeseen areas of enquiry. The reverse of this is likely to be true of most forms of interviewing (aside from completely structured interviews) which may be adapted either within an interview or between one and the next.

In addition, trainers and trainees receive many requests to complete surveys, and the responses may be hurried or peremptory, or the questionnaire may even be discarded altogether. These factors might contribute to obtaining superficial or incomplete data, and this may be due to either fatigue or resistance from potential respondents. For these reasons, I felt that this method would be unlikely to obtain the complex and subtle data I needed to address the research questions, and thus it was not pursued.

Focus Groups

The third method of enquiry I considered was the use of focus groups, as this would also create the opportunity for issues to be discussed between interviewer and respondents and would provide a more direct means of addressing assessors and trainees than questionnaires. Although such a method might allow general issues to be explored and issues defined, it would be difficult to pursue the areas further with individuals in the group situation.

Though a group discussion might encourage some participants to think and speak more widely, possibly encouraging responses, DiCicco and Bloom (2006) make the point that the process's public nature potentially prevents in depth exploration of individual views. Individuals' responses may be influenced by others in the group, leading to opinions trending toward an average, and there is also the potential for dominant figures to over represent their views. I felt that the data I was seeking would require a more individual approach and thus did not use this method.

Quantitative Methods

As well as considering various qualitative methods, I also considered whether a quantitative method might address the research questions, such as the use of questionnaires with rating scales whose data could be statistically analysed. A quantitative approach would need a large number of respondents were it to be subjected to statistical analysis, and I had no guarantee that I would receive sufficient responses in order to reach statistical significance.

In addition, opinion on the research questions is likely to be quite subtle, and thus difficult to 'score', and statistical analysis would be unlikely to reveal the nuanced nature of the data, particularly slight differences of opinion between interviewees. The latter might be obscured by statistical analysis, as might important minority opinion giving additional insight into the data. Overall, the wide range of opinion likely to be offered in response to the research questions suggests that a purely quantitative approach would be unable to gain sufficient insight into participants' views.

b. Rationale for the study sample

The key issues to be addressed initially would be which groups of potential respondents to approach, and the numbers one hoped to recruit. The opinions of practising radiologists and trainees participating in (and using the results of) WPBA would be the best sources of information upon a relatively new development in radiology training, particularly as the research questions concerned matters such as users' realisation of and influence upon the WPBA process.

It was important to gain the perspectives of both assessors (trainers) and assessees (trainees), as both the similarities and differences would be important in addressing the research questions, and the two groups might have distinct opinions on the process. Invitations to participate were sent to all consultant radiologists working in two large local teaching hospitals, and all trainees on the local training scheme eligible to join the study (see trainee exclusions below), comprising seventy nine invitees overall.

Due to ethical issues predominantly regarding my interviewing trainees (see next section), certain groups were excluded from the research, and these were as follows:

1. Any trainee whom I'd previously assessed, by any means.
2. Any trainer or trainee who had formally complained about the conduct of an assessment (none).
3. Any trainee with personal issues or difficulties which I'd had to address or help resolve in my role as Training Programme Director (very few).

In addition, very junior or very senior trainees were not invited to join the study, as they would not have had significant personal experience of undertaking WPBA in radiology, by virtue of either being too newly arrived or by having virtually completed their training prior to the assessments' introduction.

This meant that I would not gain the most senior trainees' perspective on what had gone before the introduction of WPBA, but enough fairly senior trainees (who had experience of assessment pre and post WPBA) were invited to cover this area. Aside from the excluded groups, all local radiologists with potential experience of WPBA were invited to join the study.

Consideration was given to interviewing subjects from neighbouring or even distant training schemes. This had the potential advantages of possibly increasing the overall applicability of the study findings by interviewing over a wider geographical area, and reducing the negative effects of my interviewing trainees I knew, as discussed in the ethics section. However, I decided against this suggestion, and the reasons for this are discussed below.

I wanted to gain an in-depth appreciation of the issues pertaining to a centre and its training scheme and felt that I might end up with more superficial data if I spread the interviews too thinly between various departments. Although it is possible that practice in one centre might be distinctive, WPBA has been introduced as a standardised process throughout the UK.

Though I was keen to gain a deep understanding of local issues (and maybe help solve some of them) I hoped that the data analysis would allow me to separate local from

general issues, and that some of the findings would still be likely to have some relevance to other centres. However, by basing my research on one centre and its training scheme, I would have to acknowledge that this would potentially impact upon the generalisability of the results.

In addition, the practical issue of scheduling interviews around my and others' service and training commitments proved quite challenging at times, and this challenge would have been greatly increased had I interviewed elsewhere. As a single Ed D researcher, it was important to consider practical limitations of time and scale upon what would be feasible. In view of this, consideration was given to visiting a neighbouring hospital for a day and interviewing several subjects, one after another, but this was impracticable in the face of busy schedules of clinical work.

It was also felt important to review each interview prior to the next one, in order to make iterative improvements to interviewing technique and include new issues, and this would have been impossible had I undertaken several on one day. Though there may have been some tension between this iterative progression and undertaking standardised interviews which might have potentially affected the reliability of the data collection, I was keen to refine my interview questions and technique to ensure the process was as good as it could be. Another practical argument against undertaking several consecutive interviews was that I found that interviewing required much concentration, and this would have been difficult to sustain through many closely spaced encounters.

The process of recruiting interviewees to the study was straightforward and is detailed in section f. Having decided that I would aim to interview twenty subjects, I was able to

interview the twenty who responded positively and did not have to send out any reinvestigations. I did not receive any replies stating that an individual had decided not to participate, and presumed that those who did not reply were either not interested in joining the study or, in a few cases, may not have received the invitation. Whether the study sample was representative is open to question, and the implications of self selection are discussed on the next page.

The trainers interviewed came from a range of radiological subspecialties, represented by:

- Paediatric Radiology: 3
- Nuclear Medicine: 1
- Gastrointestinal Radiology: 2
- Vascular Radiology: 2
- Chest Radiology: 1
- Oncological Radiology: 2
- Computed Tomography and Magnetic Resonance Imaging: 1

Although there was a theoretical possibility that the trainers might not be assessors, and thus just users of the assessment data, rather than undertaking WPBA themselves, in practice all the trainers were also assessors. The trainees covered the full range of the year cohorts I had approached, with a slight preponderance of those who were more senior. Further demographic information on the interviewees (including self-estimated previous experience of WPBA) is included in appendix e.

Although the sample might seem fairly representative of the assessors and trainees based upon the data above, there were ways in which it was unrepresentative of all those invited. Although the invitations were wide ranging, those who responded positively were likely to be those who were interested and engaged in training or who might have a personal or other reason for wanting to help me.

From my own prior knowledge of my colleagues, it was clear that the respondents did not include disengaged trainers or poorly performing trainees, and all of the trainees who responded could be regarded as enthusiasts who were doing well. As participation in the study was voluntary it is difficult to see how this could be avoided, but it is important to acknowledge that some groups were not included and thus important negative views may not have been represented.

It would have perhaps been possible to address this by sending reminders to those who did not respond to the first invitation, in the hope that some from the latter groups might have joined the study, but I felt that this could be seen as trying to manipulate the sample. In addition, describing colleagues as disengaged or as poor performers is inevitably a subjective judgement, and I did not think it was appropriate that I should skew the sample according to my own views. A second invitation might as easily have been ignored as the first, and if any poor performers did respond, there would be a strong possibility that I would have been personally involved in trying to resolve such trainees' issues, making them ineligible to join the study.

It would have been inappropriate to use any inducement (e.g. voucher or payment) or other means to encourage the participation of these groups, as this would have meant my

extraneously manipulating the sample and the incentive might have affected the data offered by the interviewees so recruited. Thus I took the decision to interview only those who had volunteered, whilst accepting that this would have to be acknowledged when the data was analysed. An example of this is that one would have to acknowledge the possibility of the existence of contrary or negative views in comparison to those voiced by the interviewees, and accept that a seemingly individual or minority view might be more commonly held by those who weren't interviewed. The representativeness of the sample is reviewed further in section f and chapter ten.

c. Ethical issues

Ethical issues which might have impacted upon the study included obtaining voluntary informed consent, ensuring participants knew they had the right to withdraw at any time, ensuring no harm came to interviewees, and ensuring a protocol was in place to enable disclosure of any unacceptable practices (such as allegations of bullying) were any revealed during the interviews. The study consent form and participant information sheet (included as appendices) were designed according to National Research Ethics Service (NRES) guidelines (National Research Ethics Service, 2009) to ensure these issues were addressed.

An additional issue potentially impacting upon this study was my own position as both consultant radiologist and director of the radiology training programme. Although it is possible that issues regarding authority relationships might affect interviews I undertook with my consultant colleagues, these would be potentially far greater in those with trainees for whom I had some managerial responsibility. Despite the exclusions noted in

the previous section, the question of authority relations remained pertinent with regard to the interviews.

These included trainees feeling coerced to join the study, and when they were being interviewed, feeling that they might not want to offend, or might want to please, an interviewer who had some managerial responsibility for them. In order to try and mitigate both these issues and those highlighted in the first paragraph, the study proposal followed The British Educational Research Association (BERA) revised ethical guidelines for educational research (2004), but despite this, the influence of the interviewer/interviewee relationship is unknown.

The issue could have been addressed by an alternative individual conducting the interviews instead of me, such as another consultant radiologist involved in training or even one of the administrative staff. However, the same issues with regard to power relations between interviewer and trainee interviewees and trainees possibly giving answers they thought would please the interviewer might still have occurred if a consultant colleague had interviewed, and it is unlikely that they would have had the time available to interview.

It would have been difficult to find an administrator with the specialist knowledge of the subject required to probe specific areas pertaining to WPBA in radiology, and even if such an individual were found, interviewees would have been aware of, and possibly influenced by, my presence behind the scenes. In addition, I felt it was important to participate as fully as possible in the research project for my degree, and would have been reluctant to have been involved second hand.

Having considered these factors I decided to conduct the interviews myself. I would have to acknowledge my position and any issues it might raise in the study write up, but felt that the advantages of my interviewing outweighed the disadvantages. Interviewing one's own trainees is one of the disadvantages, and although interviewing trainees from another centre might provide a solution, the practical difficulties in undertaking this (outlined in the previous section) meant that it was not pursued.

Previous experience of interviewing trainees had shown me that they were not afraid to criticise initiatives which I had helped to develop (Ramsden, 2009) and the subject of the current study (development of WPBA in radiology) was not something in which I'd been personally involved.

As intimated above, a potential source of bias with regard to my interviewing was my personal standpoint with regard to WPBA, and how my conduct of the study might lead to deriving conclusions which might develop or support my position. I view myself as a sceptic with regard to WPBA, which would undoubtedly influence how I collected the data, which issues I would pursue, and my interpretation of them.

Although I'd played no part in the development of WPBA, it was part of my role to institute it on our radiology training scheme, and thus interviewees might presume I had a vested interest in wishing it to succeed, influencing their view of my standpoint as a result. The interviewees were not told of any personal view I had regarding WPBA, but might have assumed that due to my position I was an enthusiast, and tailored their answers accordingly.

Factors such as personal scepticism regarding WPBA, yet possibly being viewed as an enthusiast after introducing the scheme locally show how my actual or perceived standpoint might influence both data collection and analysis. An example of how this could affect data collection might occur during an interview, where I and an interviewee (both with individual standpoints) might work together, even subconsciously, to build a version of their views which suited us both. Although this represented a potential risk to data collection, I cannot cite an example which occurred.

Removing the influence of my standpoint in this process is very difficult, though it might be mitigated by discussing the data and its interpretation with my supervisors, and inviting radiological colleagues to review the study and its findings, to see if they feel it reads like a true reflection of the issues. However, I do not think it is possible to remove my personal standpoint completely, and should acknowledge this.

d. Obtaining ethical approval

The study proposal was submitted for NHS Ethical review in August 2010, and I subsequently attended the review meeting in October.

I was asked why I proposed to exclude those trainees who had complained about the conduct of any assessment, as interviewing them might yield useful data. I explained that a complaint would be very unlikely, and although I'd be happy to interview any respondent who was dissatisfied without formally complaining, I felt it was too intrusive to interview those with a formal grievance, particularly as I might be involved in trying

to resolve it. The committee accepted this and to date no local trainee (to my knowledge) has complained formally about an assessment.

The committee also asked if assessors and trainers were the same people. I explained that they frequently were, and although it was possible that some trainers might use the results whilst not undertaking assessments themselves, this was very unlikely to actually occur. I felt it was legitimate to invite both groups to join the study, as I felt that users' views could be important in my evaluation of WPBA. The committee accepted this, and as mentioned previously, the twelve trainers who responded and were subsequently interviewed were all assessors too.

Finally the committee asked that the participant information sheet should include detail of how any allegations of bullying or other untoward events would be dealt with, and I was able to amend this to their satisfaction, ensuring that the document met the NRES standards mentioned in the previous section.

The process felt straightforward and submission of the explanations and amendment led to a favourable ethical opinion for the study proposal, and research and development approval for the project to proceed was given by The Leeds Teaching Hospitals NHS Trust in November, 2010.

e. Trialling of the interview schedule

The semi structured interview schedule which was used in the study was initially compiled from my personal experience of WPBA, aiming to address the issues raised by the research questions. The main headings were discussed and then amended following

comments from my supervisors and three volunteers (one trainer, two trainees) with whom I undertook trial interviews, and these encounters also helped me to improve the schedule in detail (see below).

Of the trial subjects interviewed, the first trainee had been appointed to a consultant post just prior to the interview, the second had completed just over three years of training (out of five) at the time of the interview, whilst the trainer was a consultant colleague in a teaching hospital. The interviews were conducted and recorded without incident, and subsequently anonymised during transcription. The subjects also provided feedback upon the schedule used in their interviews.

The first interviewee commented upon the use of jargon in the interview schedule, and when it was revised, words such as formative and summative were replaced with terms in everyday use, such as developmental and final. He also suggested that I should add enquiry regarding certain practical issues pertaining to WPBAs, such as whether trainees would selectively approach assessors perceived to be generous, and if they would discard the results of poor assessments.

The first interviewee also said that he perceived the assessments as liable to be influenced by the assessor's personal view of the trainee, and he felt that it would be in doctors' nature to try and pass them with high marks, rather than regarding lower scores as a developmental opportunity. It proved difficult to unpick some of the issues related to the assessment of complex clinical activity, and discussion of the behaviours associated with it tended to focus on predictable events, such as the pre procedure team briefing. Finally, the interview tended to become a little repetitive at its end.

The second interview was conducted using an expanded schedule incorporating the first interviewee's comments and my own improvements. I included enquiry about trainees seeking assessors perceived as generous, and the effect of WPBA upon future learning and professional relationships, none of which featured in the first interview. As a result the second interview took longer, but there was less repetition at its end.

The second interviewee noted the difficulty in adopting generic checklists (developed by other specialties) for use in WPBAs in radiology and wondered if they were appropriate for the assessment of complex work. He felt such lists were good for predictable events, but less useful for assessing all the steps inherent in more complex procedures.

He felt it was important I clarify with interviewees which of the assessment tools I was particularly interested in (and ask them how many they had done) when discussing them during the interviews. These questions were subsequently incorporated into the interview schedule, and I was careful to clarify which assessments I wanted to discuss when undertaking the interviews.

An area he mentioned which I had not considered before was trainees wanting observed episodes which had gone well to be included retrospectively as WPBAs. As a result, enquiry about retrospective assessments was incorporated into the next revision of the interview schedule. Both this and the temptation to discard results perceived as poor fed back upon the first interviewee's point about trainees regarding the assessments as pass/fail tests, rather than a developmental exercise.

Aside from these different issues discussed by the two trainee interviewees, there were also some areas of agreement between them, examples of this being that both thought

WPBA in radiology was better suited to formative than summative assessment, and both viewing the assessments as a combination of a test and a developmental opportunity.

Both interviewees also thought that the assessments felt more like a performance or special effort on the part of the trainee, rather than an authentic view of their routine practice, and this might explain why some trainees might discard the results of episodes they perceived had gone badly.

When comparing the two trainee trial interviews, there was a difference between the two interviewees' responses, with those of the second being more guarded than the first. This could be explained by the different seniority of the two, as the first interviewee had finished training, whilst the second had nearly two years left to complete.

The second interviewee might have been taking care not to cause any offence, possibly illustrating the effect of the authority relationship between him and I, although I did not ask him about this explicitly. If this was a contributing factor to his responses, it reinforced the point (made earlier) that I should be alert to the likelihood that trainees' answers could be influenced by my position, and take account of this when analysing the data.

The trial interview with the trainer took place some time after those I had undertaken with the trainees, and used the schedule developed after the second trainee interview. It was clear that there were areas where the trainer had similar views to the trainees, such as seeing WPBA as predominantly formative in nature, but the interview also raised issues which were not discussed in those with the trainees.

An important example of this was his feeling that he knew whether trainees were progressing satisfactorily or not overall, and that this allowed him to make appropriate allowances for trainees who performed surprisingly (to him) well or badly during an episode of WPBA, raising questions of where and when the actual judgement was made.

In order to pursue this suggestion that assessors' views might be influenced by their prior impression of a trainee, I ensured that enquiry about this was added to the interview schedule. An additional inference from this was the suggestion of the continuing primacy of consultant opinion in evaluating trainees despite the introduction of formal WPBA, and I ensured I discussed this in the subsequent (non trial) interviews.

Aside from amending the interview schedule with respect to the issues he raised, this interview played an important part in changing my interviewing style. Review of the transcript revealed that I was asking a number of closed or leading questions, thereby restricting or directing the interviewees' responses. As a result I tried to ensure that when I raised issues during the study, my enquiries were phrased in an open and non directive manner, as prompts which allowed respondents free rein to state their views, without asking them closed questions. I hoped that by doing this it would mean issues were not missed, whilst allowing interviewees to comment without restriction.

In terms of assessing whether the interviews addressed my research questions satisfactorily, I felt that the results were mixed. It was easier to directly address the questions regarding users' realisation of WPBA and their judgement of the assessments' utility, rather than that which enquired about how they influenced the assessments.

However, general discussion regarding the conduct of the assessments meant that I was

able to acquire much useful data regarding users' influence upon WPBA, though perhaps indirectly. The trial also demonstrated the need for me to take opportunities to deepen the discussion when subjects arose which would give greater insight into the conduct of the assessments.

Thus the trial interviews were useful as they helped me develop the interview schedule by removing jargon and introducing issues (e.g. approaching assessors perceived as kind, retrospective assessment requests) which I had not considered before. In addition, analysis of the interview transcripts with my supervisors suggested I had adopted a closed and directive style of enquiry at times, and that this should be replaced by more open questioning.

The importance of the latter was emphasised by Hunt et al (2011) in a paper discussing the transition healthcare professionals should make from clinical interviewing to the techniques appropriate to qualitative research. The authors contrast the often narrow and focused nature of a clinically led interview with the wider and more open style of questioning needed in a qualitative research interview, in which the interviewee is regarded as an expert revealing the nature of their experiences. As a result of both this and my supervisors' comments, I sought to achieve a broad and open style of questioning during the interviews, examples of this being to ask interviewees to amplify points they found interesting, or asking them to try and think of an example when discussing an issue they thought significant.

Because of the wider scope of enquiry and altered interviewing style developed as a result of the trial interviews, I chose not to include (the likely more limited) data

obtained from them in the analysis. Additionally, I was subsequently able to recruit twenty interviewees, which represented the number I'd hoped to include in the study, so the trial interview data were not required in order to reach my numeric target.

f. The method in practice

As discussed in the previous section, the interview schedule was modified according to the results of the trial and my supervisors' suggestions. It then formed the basis of all the semi structured interviews I undertook, and is included as appendix b. However, as the study progressed, it became clear that parts of the interview schedule elicited more material from interviewees than others, and the consequences of this are addressed in chapter ten.

Once I had obtained ethical approval for the study, I wrote to all eligible radiology trainers and trainees (forty eight trainers, thirty one trainees) in Leeds, inviting them to participate in the study. The invitation letter was distributed as an e-mail attachment, and was accompanied by a copy of the participant information sheet, included as appendix c. I hoped to interview twenty subjects, and those interested in participating in the study were invited to contact me by e-mail.

By this means I received twenty positive responses, twelve from trainers and eight from trainees, rates of twenty five and twenty six percent respectively. All who responded joined the study and were subsequently interviewed, and although I had been authorised at ethical review to send non respondents a single reminder, this was not pursued (see below).

I'd originally hoped to recruit equal numbers of trainers and trainees within my sample, but of those who responded, sixty per cent were trainers. This was reflected in the greater numbers of trainers eligible to join the study, and the proportions of positive responses to my invitation were almost equal between the two groups. In view of this I took the decision not to try and redress the balance by inviting more trainees to join the study, as it seemed inappropriate to intervene (see section b), and I interviewed all who had responded freely to my initial invitation without the pressure of any further contact.

The interviewees comprised thirteen males and seven females, roughly paralleling the gender split of the invitees as a whole (fifty five males and twenty four females). Further demographic information, including self-reported numbers of assessments undertaken (range 1 – 50), are included in appendix e. All who joined the study gave written consent to be interviewed (consent form included as appendix d), and each subject was interviewed once, the first on March 29th and the last on November 30th, 2011.

The interviews took between forty six and seventy three minutes each and I was able to resolve minor practical difficulties such as finding mutually convenient times for interviews and private places in which to conduct them without disturbance. In order to avoid exacerbating issues relating to my position with respect to the interviewees, all interviews were either undertaken in their offices or on neutral ground.

All of the interviews were recorded and professionally transcribed. Care was taken to space the interviews out in order that none was undertaken prior to my reading the transcript of that which preceded it. By this means, new issues or emphases arising in one interview, could, if appropriate, be pursued in the next. As a result, the interview

process was refined through the study, though the basic structure and semi structured schedule remained little changed throughout.

g. Analysis of the data

After they had been typed, the transcripts were securely stored in a locked filing cabinet. During my initial reading of each transcript, I rendered every interview anonymous and corrected any misspellings. I also took the opportunity to write a précis of each interview, making it easier to recognise key issues which had arisen. These summaries served as guides to the transcripts, making them easier to analyse.

The interviews were later electronically coded using Nvivo 8. I developed a system of headings and subheadings, anticipating the main themes I expected to arise from the data, with particular reference to those which would address the research questions. The system was largely derived from my personal experience of the interviews, and by reading the first two interviews, with minor additions being made as I coded the remainder of the transcripts. The complete coding system is included as appendix k.

The transcripts were then individually imported into the software and blocks of interview text assigned to a heading or subheading. Assigning text was a means of dividing the data into both broad and narrow areas, ready for thematic analysis. The decision to utilise this means of analysis is discussed in the next chapter. It also had the effect of expanding the data, as a passage of text might be relevant to more than one heading, and thus assigned more than once.

After all twenty transcripts had been entered, I went back and recoded the first two, in order that all of the transcripts had been analysed when I had had some experience of using the coding system, in an attempt to improve consistency with those coded later. This raised surprisingly few issues, as the recoding of the first two interviews led to very similar results to those which had been obtained originally, aside from diversion of a little data into sub codes which I'd added during the overall coding process (see above), hopefully slightly improving the analysis of these interviews. Data entered under headings and subheadings were grouped into broader sections for analysis (see below).

The coding system's headings and subheadings were partly based upon the original questions in the interview schedule, and partly upon additional issues which arose during the interviews, such as responsibility for case selection. By doing this, it was hoped that many of the interviewees' responses could be easily mapped to appropriate headings and subheadings.

However, it was also important to take account of the broader picture, and as a result the headings and subheadings were grouped into wider sections, such as attitudes, assessment of complex work, relationships, reflection and feedback, and long term effects. There was inevitably some overlap between sections and headings, meaning that some issues might appear under more than one, but this was preferable to issues falling between categories.

The coding system as presented in appendix k demonstrates a very large number of headings and subheadings, but this ensured that no data were lost. As stated above, the disadvantage of this approach was that the same data might be assigned to multiple

codes, leading to duplication and difficulties managing the large amount of multiplied data.

The initial coding system also included a miscellaneous category, in order that data which did not fit with any of the headings were not lost to the study. This might arise if an interviewee answered a question from a new perspective or introduced a new issue during the interview. The material in this category was subsequently classified under new headings, developed retrospectively, which meant it could be included in the overall data analysis.

Development of the themes was firstly driven by the frequency with which certain issues were raised by interviewees, indicating their importance to respondents. This contributed much to the realisation in day to day practice theme, as what occurred routinely when trainers and trainees engaged in WPBA was frequently raised at interview by respondents. The second factor which guided the development of the themes was an attempt to address the research questions.

The assignment of material to themes led on to the process of data analysis, during which the whole range of respondents' views with regard to any particular issue could be interrogated. This process helped to highlight areas of consensus amongst interviewees, in order that issues over which there was wide agreement might be reported, and the underlying reasons analysed.

Areas of disagreement could be similarly analysed, and the reasons for non consensus explored. This was particularly pertinent when comparing and contrasting the views of assessors and trainees and where there were definite differences or particular sources of

agreement either between the two groups (or within groups); this was discussed within the relevant theme.

Even allowing for duplication, the amount of material generated by the coding system was considerable. I read this material, writing a précis of each interviewee's views, ensuring that they were correctly categorised as either a trainee or trainer. I then made a précis of each group's view within the major themes, with the aim of identifying majority or consensus opinion, finding areas of disagreement, and highlighting interesting minority or individual views. This was the method I used to analyse the data, and compare the views of trainers and trainees.

Inevitably, my interpretation of the data would be personal, ranging from the coding system, through which parts of the interviews I felt were important, to my individual interpretation of the material and its messages. In order to try and mitigate this, a sample of the interview data was also given to a colleague with an educational interest, and she was asked to independently interpret it. I did not impose any constraints upon this process, and she was not asked to use any pre determined system or seek any particular findings when interpreting the data.

Allowing a second individual to interpret the data, and then compare her interpretation with mine, would hopefully lessen the effect of one reader's potentially idiosyncratic view. Once we had both interpreted the data, we compared our written notes to look for both areas of agreement and those where our views were divergent.

Generally there was a good level of congruence between our interpretations of the data, particularly major themes, with both of us able to categorise, for example, material

regarding interviewees' views on the purpose of WPBA beneath similar headings.

Despite occasional disagreement regarding minority or outlying views, this process increased the overall level of confidence that the data had not been subject to an unusual or idiosyncratic view.

4: Introduction to Data Analysis

a. Selection of data analysis method used

Various methods of data analysis were considered for use in the study, including critical incident reporting, case studies and thematic analysis. Each method is evaluated in turn, prior to a longer discussion of the means selected (thematic analysis).

Critical Incident Reporting

Critical incident reporting would have meant my analysing significant events reported by interviewees, and using these as a framework of important occurrences upon which to base my findings. Such reporting has the advantage of highlighting events which may have a strong bearing upon the key issues for the study, although what is deemed critical would inevitably be subject to the judgement of the interviewer, and interviewee providing the data.

This approach has the possible disadvantage of obscuring routine practice (in which I was interested) by highlighting unusual or idiosyncratic occurrences. It has the further potential disadvantage that some interviewees may be able to recall critical incidents with ease, whilst others struggle to do so, leading to over and under representation of some individuals in the analysis.

It must be admitted that the interview schedule did not include requests for interviewees to detail critical incidents, such as individuals' experience of very good or poor practice, and I did not seek to draw them out as a matter of course. Despite this, some

interviewees did report occurrences which might be termed critical incidents, these tending to appear during general questioning, whilst others did not mention any at all.

As I felt that it was important that the study data were representative of all participants' views and that there was a need to view day-to-day practice (as opposed to unusual events), I decided against using this means of analysis, and the format of the interviews reflected this. However, it was still possible to include significant events or minority views within the thematic analysis as a whole, and where possible, I did this.

Case Studies

Another means of analysing the data would have been to have selected particularly interesting or illustrative interviews as case studies and scrutinised these in depth. This would have provided very detailed information from particular interviews which might otherwise have been obscured by a more general system of data analysis, allowing individuals' views (and the reasons they held them) to be critically viewed, with the aim of discovering the greatest possible insight into WPBA.

Although case studies and thematic analysis are not mutually exclusive, the potential disadvantages of this means of analysis are similar to critical incident reporting. The choice of which interview might be illustrative would be subject to my opinion and some interviewees' data might not be included. As noted above, I was particularly keen that all the interviewees' views could be represented, and for this reason, I did not use this means of data analysis.

As with interviewees' reports of critical incidents, I felt that it would be possible to highlight cases of particular interest as part of the thematic analysis, and thus elements which might have appeared as parts of case studies would not be lost. In theory, one might utilise both forms of analysis (case studies and thematic analysis) but to fully engage in both would not be practical in terms of remaining within the word count allowed for the thesis.

Thematic Analysis

Thematic analysis concerns the reading of the interview transcripts to look for themes of particular interest emerging from the data, and deriving areas of both consensus and disagreement cited by interviewees. The main reasons I chose thematic analysis of the data for this study were the potential for all interviewees' views to be included and to identify commonly held ideas and experience. Despite thematic analysis's role in illustrating areas of consensus, it may also highlight minority or differing points of view, ensuring the representation of all shades of opinion.

Overall, it was felt that this would represent the best means of viewing day-to-day practice, and would allow areas of interest raised by interviewees to form themes which had not been considered previously. Some of the themes used did emerge from reading the data, but others (such as the purpose and utility of WPBA) represented pre selected areas of interest in order to address the research questions. Thematic analysis has the further advantage of allowing elements of the other means of analysis to be incorporated within it, rather than excluding them.

I felt that a means of data analysis which treated all the interviews equally (all of the other methods discussed do lead to some interviews being accorded greater prominence than others) and potentially included material from all those interviewed would be most likely to be representative of day-to-day practice.

It would be disingenuous to state that all the interviewees' issues emerged at random rather than as the result of some directed questioning, and how each theme was derived is discussed in the next section. The four themes which were finally selected broadly formed an empirical response to the research questions.

Although the data were grouped into sections, headings and subheadings during coding, the themes used during data analysis were not directly based upon individual codes or subheadings, and thus cut across coding boundaries to include data from different sections. This allowed data from different areas to be integrated into over riding themes, drawing material together to address the research questions. An example of this is the combination of data from the attitudes (formative versus summative purpose of assessments) and long term effects (dealing with poor performers) sections within the theme of WPBA's purpose.

b. Choice of Themes

Once I had decided to utilise thematic analysis of the data, it became important to select themes which would both reflect the interview data and allow me to address the research questions. Although some themes arose (one and four) because they were areas of interest I was keen to pursue, and data were directly sought in the interviews to address

them, others (two and three) arose because they represented areas which interviewees were keen to discuss, whether or not they had been led to them by the interviewer.

Thus, not all of the themes chosen were the subject of direct questioning, but some might concern issues which arose indirectly or during subsequent discussion, generating much material to analyse. By basing my choice of some themes upon material which arose during the interviews, I aimed to analyse data concerning issues important to respondents. This represented a mixed means of deriving themes, and it would be important to acknowledge that deciding which material was important to respondents (aside from basing it upon the volume or frequency of discussion during the interviews) was subject to my judgement.

The first theme chosen was users' views regarding the purpose of WPBA, as this represented an important issue to both the study and interviewees, judging by their readiness to respond to direct questioning and the amount of data generated. This could be regarded as a pre selected theme which served to illustrate a facet of the interviewees' underlying attitude to WPBA. Respondents' views upon the assessments' purpose also appeared to potentially influence their opinions upon other areas such as the day to day realisation of the assessments and interviewees' influence upon the process (research questions 1 and 2).

By contrast to the first theme, the second, concerning the realisation of WPBA in practice, arose as a consequence of answers provided by interviewees in response to many questions, which enabled me to gain an overview of their opinions of how the assessments were enacted day to day. That there was a large volume of data available on

this topic is perhaps not surprising, as all interviewees were active participants in WPBA. Thus rather than it being pre selected, the choice of this theme represented an area important to respondents which they were keen to discuss, and provided an example of how material from different parts of the interview might be combined into one area for analysis.

The third theme concerned the influence users might bring to bear upon the process, and how this might affect the enactment of the assessments. The origin of this theme was similar to the second, being derived from the analysis of a large number of responses during the interviews. Again, data were derived from different parts of the interview, and in comparison to the other themes, more material in this category was derived from indirect questioning and discussion, rather than direct enquiry. Of the themes selected, the second and third appeared the most inter related, and some overlap between the material included in each resulted.

The fourth theme included interviewees' opinions regarding the utility of WPBA, their views upon areas of strength and weakness, and suggestions they might have to improve the process. As with the first theme, this could be regarded as a pre selected area of interest within the study, with data being obtained by interviewees being led to this area. However, as with all of the other themes, it was an area which interviewees seemed keen to address, providing a large amount of data for analysis.

Thus the four themes under which the data were organised are summarised as follows:

1. The purpose of WPBA

2. How WPBA is realised in day to day practice
3. Users' views regarding how they influence the realisation of WPBA (including comparison with central policy regarding its recommended enactment)
4. Users' opinions of the utility of WPBA, its weaknesses and how they feel the process might be improved

Within the reporting of the analysis, the respondents' views are represented by direct quotations from their interviews. Each quotation is attributed to a specific interviewee (without compromising their anonymity) by referring to them as trainee one to eight, or trainer one to twelve.

5: Theme 1:

The purpose of workplace based assessment

This chapter briefly outlines central policy regarding the purpose of WPBA, before reviewing opinions of the assessments' purpose expressed by trainees, followed by those of trainers. I was interested in interviewees' views of WPBA's purpose as I felt this was likely to influence how they approached and utilised the assessments. To develop this, I sought further opinion regarding WPBA's purpose by asking interviewees whether they felt assessments could be passed or failed.

This is followed by the presentation of some issues regarding the assessments' purpose raised by interviewees, including some trainers' use of WPBA to identify poorly performing trainees, and the influence upon the purpose of numeric targets, prior to a short summary at the end of the chapter.

Much of central policy regarding WPBA was discussed in chapter 1. It was introduced into radiology in autumn 2010 as a standardised process described in the RCR's curriculum (beneath the heading Formative Assessment) as; '*the cornerstone of assessment for day-to-day practice*' (Royal College of Radiologists, 2010).

The report of an earlier trial organised by the RCR recommended that assessments of a clinical episode be used formatively, but also stated that, '*trends from multiple encounters may be used to inform summative decisions*' (Royal College of Radiologists, 2010), perhaps leading to a slightly mixed message. The following section describes

how interviewees perceived the purpose of the assessments in practice, and contrasts the views of trainers and trainees.

a. Interviewees' views of the overall purpose of WPBA

The majority of trainees perceived WPBA as a developmental (the more common term interviewees used) or formative exercise, examples of their views being demonstrated by the following quotations;

'Progress, development of the trainee, the basic skills and knowledge of a trainee' (Trainee 8)

and

'A teaching tool. At the end of the day that's what I believe it is. It's an assessment tool. I learn from my mistakes and it documents the progress that I'm making.' (Trainee 2)

Many trainees saw provision of feedback as an integral part of the process, allowing areas of weakness to be highlighted and improved, as noted in the second quotation. The need to know where improvement was required was highlighted by a trainee, who said;

'I think the role of the assessment was to allow you to sit down and reflect on how you performed on a case yourself, and then to see how that compared to what the assessor felt about you, and it was also good to see which bits they thought you'd done well and what you needed to improve on.' (Trainee 3)

A stronger view was that the assessments obliged the trainers to take greater account of trainees' needs, ensuring that time was found to deliver feedback. Thus in addition to the assessments' primary purpose, some trainees identified what they felt was another

purpose which was that the scheme ensured that assessments occurred and that feedback was delivered, the same trainee adding;

'Actually having a scheme in place that means that you do have to do the assessments, if that makes sense. So before you had to do workplace based assessments then there was no drive for people necessarily being as interested in how you were performing in each case and whether you were getting better.'
(Trainee 3)

This quotation illustrates how the mandatory nature of WPBA can impact upon the relationship between trainer and trainee, and make formative assessment a sustained process through a period of training. As well as helping to ensure the delivery of feedback, some trainees intended to utilise the assessments as a means of structuring and keeping a record of their training, as evidenced by a trainee who described the purpose of WPBA to;

'form a record of what you've done, and the areas that you've gained experience in, but also to demonstrate and feedback to you how competent you are in those areas that you've been assessed in.' (Trainee 5)

This quotation suggests this trainee's intention to formalise their training and how the undertaking of WPBAs imposes a more defined framework upon both trainer and trainee. This was amplified by the trainee, who found this aspect particularly useful, saying;

'I think probably getting back to what I said before about structuring training. I think that's particularly useful, allowing you to monitor your progress, and also giving you some proper feedback time with a consultant.' (Trainee 5)

Other trainees concurred with the view that WPBA served to formalise feedback, but a minority did not see it as an advance upon what had gone before, and one trainee compared WPBA with previous practice by stating;

'I think they're there to formalise feedback from consultants. And if I think I was in a training scheme where I felt I wasn't really getting much feedback normally from my consultant trainers, then I think this would be a really valuable thing... But, as I say, in my experience, the vast majority of consultants that I've been with routinely do take that time to give me plenty of feedback already.'
(Trainee 4)

Other trainees took the view that WPBA did represent an improvement in the evaluation of day to day work, particularly with regard to knowing whether specific skills had been developed;

'I think one of the issues with assessment was that earlier, before... there's nothing to assess that we can do certain things. And these workplace assessments assess that. And I think until these WPBAs came up, there was nothing that would actually assess that you can actually do a certain procedure or interpret a certain image.' (Trainee 8)

A minority, including the respondent quoted above, felt WPBA was intended to test specific competencies, another trainee supporting this view by saying the assessments' role was to;

'demonstrate and feedback to you how competent you are in those areas that you've been assessed in, so that you can show a spectrum of activities in training, but also give some evidence to how competent you are in those areas.'
(Trainee 5)

Although not an additional purpose of the assessments themselves, one trainee was even more enthusiastic about WPBA demonstrating competencies, feeling that the results would be useful to demonstrate their abilities to third parties (perhaps future colleagues or supervisors), stating that the role of the assessments was as follows;

'I think it's twofold. I think it's to inform the trainee of what your ability is like and what your areas of improvement need to be, but it's also to inform a third party who doesn't see you at a regular stage what level you're at as well.'
(Trainee 7)

Some trainers also saw formalising the assessment process as a purpose of WPBA, but commented regarding the recording of all facets of an individual procedure within a specific framework, rather than structuring training overall (as per the trainee quoted previously), one saying;

'These workplace based assessments have to become more formal and rigid and if you're doing a complex vascular case where you have to go and discuss it with the patient first that might not be very practical unless you have a checklist which says did you go and discuss it with the patient.' (Trainer 12)

Generally, trainers mentioned a formative purpose less than the trainees, although there was general acknowledgement that the process was developmental, and should be used to deliver feedback to help trainees improve and to direct their training. An example of this approach was expressed by this trainer, echoing others' opinions, by saying;

'It's primarily for me a method by which I can feed back to my trainee each week, every day hopefully. Feed back to my trainee, how did that go, how can we do that better, and next time we do this next week we'll do it this way,'
(Trainer 4)

Despite this, the view that the assessments' purpose (and how they were designed to be used) was to demonstrate trainees' level of competence was more widely expressed by trainers, one stating;

'Their prime function is about learning and becoming competent and I think mostly we learn and become competent by having some sort of test.'
(Trainer 12)

Aside from judging a trainee's level, this quotation also suggests a view (in this interviewee's opinion) that a test is a necessary stimulus to learning and becoming competent. Trainers also felt that the assessments were useful to demonstrate trainees' progress during attachments, and in common with the trainees, saw them both as a means of documenting training and also acknowledging the trainer's input, as this trainer stated;

'I think for the purpose of documentation, I think it's very good because it allows a record to be kept of the learning process of the trainee. Also, a formal recognition that I am teaching the trainee.' (Trainer 2)

Some trainers also highlighted an additional purpose of being able to assess trainees in a real environment, one stating that the role of the assessments was;

'To see how trainees are progressing in a real environment as opposed to a rather false exam condition.' (Trainer 3)

I think that this comment suggests that the trainer wanted to make episodes of WPBA as valid as possible in assessing everyday work, and the purpose of assessing trainees in a

real environment was developed further by some trainers, who expressed a wish to observe all parts of an interaction with a patient, rather than just the technical parts of a procedure, one saying;

'I'd be watching for all those things, did the registrar introduce themselves and explain the procedure and was pleasant and using an appropriate combination in dose reduction and all of the usual scenarios.' (Trainer 10)

Another trainer from a different area of radiology expressed a similar view, saying that the assessments;

'Makes you think about whether you talk to the trainees about what sorts of things to say to the patients and relatives, what questions to ask, what information to give the patient, how to give it, and all that kind of thing as well.' (Trainer 3)

These quotations all suggest that these trainers saw an important purpose in ensuring that they observed as much of the trainee's behaviour as they could during an assessment episode, thus making the assessments as holistic as possible. As with the desire expressed previously to observe trainees in a real environment, this holistic view might be related to trainers wanting to make WPBA as valid as possible, although increasing an assessment's thoroughness may not necessarily improve its validity if important areas are still omitted. Reference to the score sheet for the assessment of procedural skills (Royal College of Radiologists, 2010) which is included as appendix i, shows that the intended range of attributes to be assessed complements these trainers' views regarding observation of an entire episode.

An indirect means of seeing how interviewees perceived the assessments' overall purpose, was my asking if they thought it was possible to pass or fail WPBAs. All of the trainees thought that the assessments were not intended to be passed or failed; as such defined outcomes were not consistent with formative assessment, which if it revealed deficiencies, would highlight areas for further improvement. A view on passing or failing the assessments, and how the trainee came to hold it, following what she saw as suboptimal use of WPBA in another specialty prior to entering radiology, was expressed in this exchange;

A: 'No, that's how I used to see them, I used to see it as you needed to pass each assessment, and it's not that, it's a case of being able to think about how you've done and to discuss it and to reflect on the comments.'

Q: 'What changed your mind about workplace based assessment?'

A: 'Probably seeing them being done properly, having a bit more time to go through them, and actually having them done on a continuous basis rather than nights before appraisals are due and trying to get through ten assessments which was how I'd seen them being used.' (Trainee 3)

As well as expressing the opinion that the assessments did not represent episodes to be passed, this trainee's view that WPBA should represent a continuous stimulus to improve suggests a view that the process is formative in nature. Approximately half of the trainers agreed that the assessments were not pass/fail, with some pointing out that even mistakes judged as significant would not lead to failure, as demonstrated by this view expressed by a trainer;

'Obviously if major pathologies are missed and the patient would be potentially put at risk then that is not a good outcome, but no I don't consider them a pass/fail assessment.' (Trainer 9)

By contrast, the other half of the trainers felt that the assessments represented a test which could be passed or failed, with some stating that this purpose increased towards the end of an attachment, as demonstrated by the following quotation;

'I suspect, near the end of the attachment, it does become more of an assessment, whereas in the beginning of the attachment, I think it's more teaching.'
(Trainer 2)

One trainer saw the format of the marking system (score sheets included as appendices g and i, and discussed later) as a factor in their regarding WPBA as pass/fail, saying;

'You can pass or fail, because the system of marking actually says, is this person performing at above or below the level that they should be expected to? So if you're performing below the level that you're expected to then by inference you're failing.' (Trainer 11)

These views suggest that some trainers see WPBA as potentially having some summative purpose, depending upon what use is made of the outcome, and this is developed further in the third theme regarding how users may influence the assessments when they are realised in practice. Potentially, the varying views between some trainers regarding the purpose of WPBA (formative or summative) might lead to different types of assessments being run by different assessors, or the same assessor utilising WPBA differently depending on the trainee's seniority or time in an attachment.

A related example of a purpose of WPBA perceived by trainers, but not by trainees, was the use of the assessments as a means of identifying trainees who were underperforming, with the aim (expressed in the second quotation) of supporting their training. Two trainers expressed their views as follows;

'I think they (workplace based assessments) are only useful if they identify someone who is failing to come up to a standard that you anticipate.'

(Trainer 11)

and

'You know, you're really using this as an early-warning system for trainees that are struggling, or failing, and, in my experience, that may be 10% of trainees. So, I think, all of this exercise is really to try and identify the bottom 10%, and to try and keep them going, coax them along, get them through to the other end.'

(Trainer 1)

Conversely, only one trainer stated that the assessments' purpose was to identify trainees who were doing well, saying;

'You can also pick out those that are way ahead, which is quite good because you can help nurture those into the right directions in their career and pick up problems, I'd have thought.' (Trainer 10)

The latter is particularly consistent with the developmental purpose of WPBA advocated centrally, and has the potential practical implication of encouraging the development of particular talents or interests, in contrast to the other trainers quoted previously, who appear to see the assessments' purpose as a diagnostic method for the identification of failing trainees. These trainers may see WPBA as formative as well, if their assessments help trainees to improve, but there may be a dual purpose (i.e. possibly summative as well) in their wanting to identify poor performers, as exemplified by the use of the term 'against' in the following quotation;

'I think we've suffered in the past from maybe not having that tool, really, to gather data against trainees that were struggling.' (Trainer 1)

It should also be noted that aside from views on the purpose of the assessments stated above, some trainees perceived the purpose of their undertaking the assessments was to meet annual numeric targets, two saying;

'Because we're mandated to do X number of workplace based assessments before our end of annual review, there will be a desire to get a tick box attitude towards getting them completed' (Trainee 1)

and

'Essentially, we've been told, as part of our training, that we have to have achieved a certain number of – or completed a certain number of – assessments over the course of a year I've pretty much stuck to the minimum requirements of what I have to achieve over the last six months.' (Trainee 4)

Requiring the achievement of such targets might be seen as a method of trying to ensure a minimum level of trainee engagement which could be recorded. Similar views to those above regarding numeric targets were expressed by other trainees and trainers, and they are discussed in greater depth in the next chapter which addresses the second theme of how WPBAs occur in everyday practice.

b. Summary

This chapter has addressed how users view the purpose of WPBA, and it is important to consider the meaning which interviewees may attach to this term. Without being led by the interviewer into specifically theoretical or practical domains, representatives of both groups discussed purpose ranging from their view of central policy, to their own opinions regarding WPBA's intended usage, to very practical perceived purposes such as documenting training.

This may partly account for the heterogeneity of the responses and clearly varying aspects of WPBA were important to different respondents. It might be argued that some of the purposes cited below were actually consequences of the scheme. However, it seemed that some interviewees intended consequences such as the documentation or formalisation of training, and thus they could legitimately be viewed as a purpose in their view.

Interviewees were initially asked whether they thought WPBA represented a formative or summative process. The majority of interviewees viewed it as the former, with those who felt it might have a summative role generally confined to the trainer group. The way in which this view could influence the assessment process is discussed as part of theme three.

Other purposes attributed by users include structuring, documenting and formalising training, structuring being the imposition of an overall framework on training, whilst formalisation related both to this and to the conduct of the assessments themselves, particularly in the delivery of feedback. This is intimately related to documentation, which ensures that these episodes become a permanent part of a trainee's record of training, and this recording of performance may link with some trainers' view that WPBA should be used to identify poor performers.

Some users also mentioned additional purposes, such as allowing the demonstration of competence, undertaking assessment in a real-life environment, and the need to reach external numeric targets of assessments during training. These targets are set by the RCR and aim to promote trainee engagement and sufficient usage of WPBA. However,

it is likely that trainees' purpose in achieving numeric targets relates to ensuring they undertake sufficient to satisfy college requirements and pass their annual assessment, as exemplified by quotations in this chapter and further discussed in the next.

In terms of differing opinion between the trainer and trainee groups, some trainers assigned a summative element to episodes of WPBA, but this was not the purpose of the assessments in the trainees' view. There were some trainers who felt that the purpose of WPBA should include identifying poor performers, but this opinion was not voiced by any of the trainees.

Between trainers and trainees there exists a wide variety of perceived purposes of WPBA including its developmental usage (all trainees), identifying failing trainees, summative assessment (demonstrating competence), mandating and structuring feedback, record keeping, and increasing the authenticity of assessments. It is perhaps not surprising that such diverse opinions contribute to a complex picture of day-to-day realisation of WPBA, and this forms the subject of the next chapter.

6: Theme 2:

How WPBA is realised in day to day practice

At the start of the preceding chapter (theme 1), a reminder of central policy confirming the formative ethos of WPBA was included. Allied to this guidelines for the conduct of the assessments were also issued by the RCR (2010) and these are included as appendices h and j. They are important, as in conjunction with the College Curriculum, they provide fairly detailed information on the expected format of assessments (see below), and thus might be expected to guide how the RCR feel WPBA should be enacted during day to day practice.

In terms of day to day practice, the guidance for conducting WPBA states that patients should be aware an assessment is occurring and the trainee should be directly observed whilst being assessed. The College curriculum mandates that the assessor be trained in giving feedback and that this must be delivered after the assessment. The curriculum adds that the trainee is instructed to agree the case for assessment, its timing and who will assess them, with the caveat that assessors may initiate unscheduled assessments themselves. Aside from the conduct of the assessment, the guidance is concerned with how to fill in the assessment form and rating of the trainee's performance.

It is apparent from the above that there are some areas which the guidance does not cover, such as whether or not retrospective choosing of cases is allowable, who the trainee may choose to assess them and the spacing of assessments through the training year. There are also areas of ambiguity, particularly in the area of initiating assessments,

where the lead role assigned to trainees in this regard is accompanied by the caveat noted above. This allows trainers to undertake unscheduled assessments, which are presumably trainer initiated.

With such a wide ranging process, gaps and ambiguity in the guidance for WPBA are perhaps inevitable, as it is impossible to take account of every local circumstance or constraint which might impact upon the conduct of assessments throughout the UK. In addition to local surroundings, there will also be a large variety of personnel involved in the process, and some leeway (within broad guidelines) will have been allowed to take account of individual circumstances.

Participants will bring their own constraints and agendas to the process, and by constructing it together, will inevitably influence its realisation. This is explored later in the third theme, but serves to demonstrate the close relationship between the day to day enactment of WPBA and participants' influence upon the process.

The consequences of this are explored in this chapter and the next, initially by exploring interviewees' opinions of how WPBA is realised in everyday practice, and subsequently (in the next chapter) their views on how they influence the process. The chapters seek to investigate how undertaking WPBA felt to the participants, and why they acted as they did. The material in the current chapter is separated into sections, running chronologically through the process:

1. How trainees approached WPBA and users' actions before assessments (a-c)

2. How the assessments were enacted (including initiating an episode of WPBA and choosing cases for assessment) and how participants behaved during them (d).
3. How feedback was delivered (e).

a. How the trainees approached WPBA

In the context of the study and as discussed in this section, the trainees' approach to WPBA is taken to mean their beliefs and motivation prior to undertaking the assessments, viewed from both trainees' and trainers' perspectives. These beliefs and motives include examples of both strong and minimal trainee engagement and illustration of the various reasons trainees cited for their involvement, such as knowing their current level, gauging their progress and structuring their training. Finally, there is a demonstration of some trainees' expectations that WPBA would help build their confidence. This exploration of trainees' beliefs prior to WPBA is viewed separately from users' actions before assessments, which are discussed in section b.

Interviewees' views regarding the purpose of the assessments undoubtedly influenced their approach, and as noted in the previous chapter, for some this included the perceived need to reach the numbers required, whilst for others it represented an opportunity for personal development. Determining a number or percentage of trainee respondents in whom achieving a numeric target was a prime motivator could be difficult, as trainee interviewees might under report this to a trainer interviewer, in favour of reasons that they perceived as more educationally based for engaging with WPBA.

An example of a trainee who gave educational reasons for their engagement with the assessments, was one who described being motivated by WPBA's ability to;

'Trigger learning points, to highlight areas of further learning need and then to use that as a springboard to go off and do some more work as well as highlighting improvement points for the future.' (Trainee 1)

The trainee quoted above had undertaken a postgraduate diploma in education, which he said had helped give him an insight into WPBA's underlying ethos, which perhaps meant he was an enthusiast for the scheme. However, even after sharing this positive view, the same trainee later mentioned the need to achieve an annual numeric target, showing the importance of this factor to some trainees, even in cases where it was not the prime reason for their engagement.

In the preceding chapter regarding the perceived purposes of WPBA, other trainees also expressed the view that the assessments were useful formative episodes, but the influence of numeric targets was also pervasive, and two senior trainees, both of whom could be considered as successful (one medal winner, both now successfully completed training), mentioned targets' effect upon their engagement with WPBA. They had been in training prior to the introduction of WPBA, and felt that its introduction had done little to improve their training, one saying;

'So I've not really felt that I gain anything additionally from the very good training that I get already by doing that workplace-based assessment. So I've not really felt motivated to ask my trainers to take any extra time to complete these assessments over and above what I need to do as a minimum.' (Trainee 4)

Although such comment might be viewed partly as an evaluation of WPBA, and for this trainee, no great change to what had occurred previously; it still represented a facet of day to day practice. The trainee was candid regarding how little the scheme had done to alter his training, and both he and the trainee quoted below were progressing well without it. Consequently their engagement with WPBA was governed by undertaking the minimum number of assessments needed to reach the annual target.

The other senior trainee did not mention achieving numeric targets explicitly, but described engaging sufficiently to ensure her training record was satisfactory (the assessments are not formally graded by letter), whilst stating that WPBA had had little influence upon her training otherwise, saying;

'I haven't done anything differently since the introduction of the workplace based assessments other than making sure I get B's for my eportfolio.'
(Trainee 6)

One of the reasons WPBA was introduced was to improve the assessment of trainees' day-to-day work and ensure they received feedback upon it, but in some cases, it was clear that this was perceived as no improvement upon what had gone before and that these trainees' engagement was due to the fact that the assessments had been made a compulsory part of training.

As a result, both of these trainees sought to conform to the rules of WPBA in a minimalist fashion and undertake the assessments with a minimal amount of change in order to continue their training as before and not have their progression derailed by either failing to reach target numbers of assessments or poor results.

Individual trainees' level of engagement in the process thus varied, from near-compulsion (as described above), to the more positive emphasis detailed initially, and the subsequent level of engagement of these trainees might follow these initial patterns. It is important to note that those with limited engagement with WPBA were not necessarily poor performers, in fact the reverse in the instances quoted, where successful trainees did not find the assessments particularly useful.

A majority of the trainers held the view, when discussing the trainees' engagement, that it was primarily driven by the need to meet target numbers of WPBA. Some trainers added that they felt this approach led to a lack of reflective learning on the trainees' part and one felt that the trainees engaged in assessments as individual episodes, rather than using them as linked steps along a learning pathway, saying;

'I think the majority will see that as an individual case and close the chapter on it, and move on without necessarily taking home the more generic points that may have been made.' (Trainer 2)

A colleague, concerned about trainees' numbers-driven approach and a perceived lack of reflection on their part following assessments said;

'So I think the trainees potentially there is a real danger, actually, of them becoming sort of, "Oh, I've got my ten and I've done okay, so that's fine," rather than actually analysing their performance, reflecting on their performance.' (Trainer 1)

Despite these negative views expressed by trainers, an overall majority of the trainees identified educational factors rather than achieving a numeric target as major influences

upon their engagement with WPBA. However, as was seen with the trainees quoted earlier in this section, achieving the target number of assessments was also important to at least some of them, though the exact number may be difficult to quantify.

Other factors contributing to the level of trainee engagement included wanting to know their current level, gauging how they were progressing, structuring their training, and building confidence. These generally represented hopes which had been realised, confidence being built by encouraging feedback, illustrated by a trainee commenting;

'A senior registrar or consultant, who's got a lot of experience, has actually assessed me and saying... and has told me, "Okay, you are doing well." And that's a very good feedback. It boosts a lot of confidence, and that'll reflect on my work.' (Trainee 8)

Whilst another trainee said;

'If it's something which is demoralising and that would bring my confidence levels down then it's not an education tool anymore. So it has to boost my confidence and I should feel good about it.' (Trainee 6)

It might be argued that the role of the assessments in building confidence could be seen as part of the experience, rather than a factor influencing trainee engagement, but the hope that one's confidence would be improved, as opposed to receiving news that one was performing poorly, clearly formed part of some trainees' expectations, and thus could be seen as an influence upon how they approached the process. This has resonance with some trainers' wish to deliver a positive message which is further discussed later in this chapter.

Examples such as this serve to demonstrate how participants' beliefs and expectations regarding WPBA might influence how the assessments were enacted on a day to day basis, but it is also particularly important to address how the process was practically realised in the workplace. In the next section data is presented regarding the initiation of episodes of WPBA, their timing, choice of cases for assessment, choosing assessors and whether or not the assessors and assessees should have prior knowledge of each other.

b. Trainers' and trainees' actions as they prepared for an assessment

This section concerns what interviewees said they did prior to engaging in day to day WPBA, and how they might negotiate the parameters within which an assessment might subsequently occur. Within this section the subjects covered are initiation and timing of assessments, selecting cases for WPBA, choosing assessors, and whether to allow cases chosen in retrospect, all of which have the potential to be manipulated by participants (termed as gaming by one interviewee). In addition, the influence upon the assessments of trainers and trainees having prior knowledge of each other, authority relations between them, and the potential problems caused by poor trainer engagement are also discussed.

The first issue addressed is the initiation of assessments, and trainees generally appeared comfortable with both initiating WPBA and choosing cases, as shown by the following quotations;

'It's my training and my responsibility to keep an eye on how things are going, I've probably kept an eye on where the blanks are starting to emerge, in the assessments, and I've said to the consultant, I notice that we've got such-and-such on our list this afternoon. This is one of the ones that I haven't been assessed for. Do you mind if we do that as an assessment?' (Trainee 5)

and

'I just chose the people who were looking after that session, the consultant in charge of that session. And it tended to be the sessions that were perhaps not the busiest sessions, so I chose one from an acute session.' (Trainee 1)

Despite this there were occasional issues regarding trainees attempting to initiate assessments, and finding that trainers appeared too busy to undertake them, and these are discussed later in this section. In addition there were representatives of both groups who suggested either joint or trainer-led initiation of WPBA, and although this is discussed at greater length in theme four, it is also discussed below as representative of a facet of day to day practice.

A greater trainer input into initiating assessments was perceived by trainers as a means of ensuring trainees engaged with the process beyond achieving 'minimum numbers', that appropriate cases were selected, and that the assessments were evenly spaced through attachments, representing factors which some trainers felt might be neglected otherwise. This trainer felt that this was particularly important if trainees were seen as underperforming, saying;

'There's also, I think, a mandate for things to be trainer-driven when a trainee has any difficulties or educational targets to steer them in the correct direction.' (Trainer 5)

Thus the day to day practice advocated by this trainer could often differ from that suggested by central guidelines, and some trainers, including the one quoted above, felt that the leadership of the assessments might be manipulated by some trainees, who would choose easy, unchallenging cases in order to obtain ‘good’ as opposed to ‘useful’ feedback, as illustrated by these trainers’ responses;

‘If Workplace Based Assessment is used as it currently tends to be, there’s a tendency for people to say, “I’m going to look for things that I’m going to try and do well at and may have rehearsed and practised, and then go and get my assessment on those so that I get a good mark on it,” rather than saying, “I’m going to pick a challenging subject and this is actually going to help me become educated and show me how I can develop”.’ (Trainer 5)

and

‘Certainly there have been one or two trainees who actually I felt overall were struggling with the practical aspects of the attachment, but their assessments, their direct observational assessments will have been okay, because the cases they chose were extremely straightforward.’ (Trainer 3)

A further issue where some trainers felt that trainees could manipulate the process was in the trainee’s choice of assessor, where some felt that trainees tended to select trainers they perceived as generous, one saying;

‘I think there are a lot of biases, potentially, that the trainees will seek out the people who they want, and that they see as being favourable assessors.’ (Trainer 1)

As with the perceived selection of unchallenging cases highlighted previously, this trainer felt that trainees did this in order to obtain positive, rather than useful feedback.

Other trainers identified a separate issue which might also arise, concerning trainees selecting cases for assessment in retrospect, choosing episodes which had already gone well, as demonstrated in the following quotation;

'What you have to avoid is choosing cases that have been rehearsed or done in the past, in other words waiting to say, "Well that went rather well, I know, I'll ask for an assessment on that".' (Trainer 5)

These views were corroborated by a minority of trainees, who agreed that they sometimes chose unchallenging cases, occasionally in retrospect, with examples as follows;

'I have done a couple of easier ones, I suppose, to increase my numbers slightly. But I don't think I'm going to learn very much from the easy ones.' (Trainee 4)

and

Q: 'How do you feel about asking afterwards?'

A: 'A bit cheeky, because I think I might do a bit of cherry-picking, especially when I know I'm being assessed.' (Trainee 8)

A contrary view was expressed by another trainee who felt that engaging in the assessments ought to force trainees to confront subjects with which they weren't comfortable, suggesting an element of challenging personal boundaries when approaching WPBA in addition to achieving numeric targets. He felt that this was a positive attribute of WPBA, stating;

'We do lots of things that we're good at, because we find them easy and we enjoy doing them, whereas this, I've found, forces you to confront the things that

you're less good at, have less experience of, which I think is a useful aspect of it. (Trainee 7)

Trainer perceptions that some trainees might choose unchallenging cases appeared to provide some of their reasons for challenging the trainee-led nature of the process, although this view was not held by all trainers, with a minority feeling that trainees selected cases for the 'right' reasons. When asked if he felt trainees subverted the assessment process by choosing easy cases, one trainer responded;

'The vast majority of doctors have immense integrity, honesty, insight, and I have never experienced that unfortunate behaviour.' (Trainer 7)

When asked why he held this view, this trainer added;

'I prefer trainee-led. It encourages insight, it determines the interest of the individual, it puts the emphasis on the individual as a trainee being responsible for their actions as a doctor.' (Trainer 7)

This response suggested a less 'paternalistic' view of trainees and a view of them as responsible adult learners, but it represented an individual opinion not expressed by other trainers.

Both this trainer and some others were happy to assess trainees on cases chosen in retrospect, provided they felt the cases were appropriate. Although willingness to use cases in retrospect was definitely a minority view, it provided another example of trainers utilising the process in a way they saw fit, another trainer saying;

'There may be a lot for us to learn from a retrospective aspect or a retrospectoscope, so I don't think we should not use them.' (Trainer 6)

Contrary to this, most trainers favoured an approach to assessment involving material chosen prospectively, with some even suggesting that cases be chosen at random, as advocated by this trainer, who said;

'But if it was an understood system that is random, and the trainee and the trainer have no part in choosing that, then I think that's appropriate in terms of cognitive radiology.' (Trainer 11)

This trainer felt that this would be the only way to produce a truly fair system, which would eliminate any bias in case selection on the part of either the trainer or trainee, although he acknowledged that it would be very difficult to implement in practice. Such a method could be used for either formative or summative assessment, although there would clearly be disadvantages, such as the potential to not utilise potentially instructive teaching cases. This is further discussed in chapter eight.

Allied to this was some trainers' view that assessment cases had to be chosen in prospect in order that they could engage in viewing (and assessing) a trainee's entire interaction with a patient rather than just the end result, as exemplified by this view expressed by a trainer;

'I wasn't there when it happened and I turned that down as a DOPS because I felt I needed to be there to see how they managed the whole thing as far as contrast and everything else goes rather than just the report.' (Trainer 10)

This aligns with some trainers' aims expressed in the previous chapter of assessing trainees thoroughly in a real life environment. Some of the trainees expressed similar views, seeing the prospective choosing of cases as a means of engaging the assessor's attention throughout and ensuring they gained as much as possible from the assessment. When asked if cases for assessment should be chosen in prospect, a trainee responded;

'I think the trainer will look at you in a more critical way if they know that they're assessing you. So for example, if you are doing a procedure they'll watch you from start to finish and watch in particular how you do anything rather than just looking at the outcome, which is important.' (Trainee 7)

Thus the prospective selection of cases appeared to be favoured by the majority of interviewees, with representatives of both groups favouring engaging in the assessment of 'whole' cases, rather than just the result or outcome.

Beyond the single episodes of assessment, some trainers also felt that the timing of WPBA through attachments was important, stating that their formative nature was best suited by their being used regularly through a period of training. By this means the trainee might develop their skills and show progress over time, rather than trainees undertaking WPBA at the end of an attachment when they might score highly, but not have time to improve. There were trainers who felt that trainees had not adopted the culture of regularly spaced assessments, one saying;

'I don't think there is a culture yet of let's do this every week for the whole 12-week attachment which is, as I understand it, the way it's meant to be. It's not meant to be near the end of the attachment, it's meant to be all the way through and used as a tool of how can we improve on this each week.' (Trainer 4)

His comments suggested a view that some trainees were not fully engaged in the ongoing formative ethos of WPBA (despite all those who were interviewed claiming to see it as a formative process), and tended to undertake them at the end of clinical attachments. This was possibly driven by the realisation that a numeric target had to be met during a period of training, or a desire to score highly (as noted earlier) by being assessed when trainees felt they were at their most competent and most likely to ‘pass’ an assessment. These traits were perceived by this trainer, who said;

‘I think trainees, they’re gamers... I think everyone tends to sort of play games. But if they’re viewed as pass/fail there’s a tendency to go to nice people, easy cases and to do them late on, which is exactly... it’s the antithesis of the whole message and ethos of Workplace Based Assessment, which is to lead to development you need to do them early, you need to do cases that will challenge you and identify areas for development.’ (Trainer 5)

Inevitably, trainees would be unlikely to admit to manipulating the WPBA process to their own advantage (i.e. gaming) in this fashion to an interviewer with responsibility for training, but hints of selectivity in which trainers trainees would approach did arise, this trainee stating that he generally requested assessments from trainers with whom he was familiar;

Q: ‘Is there any other reason you ask familiar people?’

A: ‘I suppose you’re hoping that you get a good assessment, that’s only natural. But then I would hope that people would be honest in their appraisal. I think most... I would never expect anyone to be dishonest in their appraisal.’ (Trainee 7)

Although there were examples of trainees wishing to know the assessor (as above), and possibly gain an advantage from this, there were also trainers who wished to have prior knowledge of the trainee prior to undertaking WPBA, as illustrated by these examples from trainers' interviews;

A: 'If I have minimal experience of a trainee they come to me with a case that they've excelled on and they're generally a very poor candidate, and that is misrepresentative of the trainee's ability.'

Q: 'So what do you do when that happens?'

A: 'If I don't feel I have enough experience and interaction with the registrar I'd say let's mature our relationship and take a case in a few weeks time.' (Trainer 7)

and

'When there are assessments that have been done and have gone in the wrong direction for the trainee, that the next set or two sets of trainers are made aware of this so that they can try to remedy some of the issues.' (Trainer 11)

These exchanges did not represent isolated opinions and represent a departure from the idea that trainees might select any trainer they wished to ask for WPBA in a spontaneous fashion. This militates against the premise that individual assessments are meant to operate as stand alone episodes, allowing trainees to obtain assessments from a wide range of assessors as each WPBA is opened separately, rather than being part of a predetermined series.

Although the second trainer quoted wanted prior knowledge of trainees in order to try and help them, it still suggests that some assessors' approach to WPBA may be influenced by factors in place prior to the assessment, and a trainee's previous

performance might influence them during a forthcoming assessment. This view regarding assessors' prior knowledge of trainees is supported by other trainers, who said;

'I don't think it's as objective as it likes to think it is. I try and be objective but I'm afraid... not afraid but I think when you come along to any assessment and you know the person outside the sphere of that five minute assessment you are biased, so you do bring along things from outside the assessment, that I think I would come into it with a pre-formed opinion in many ways.' (Trainer 9)

and

'I think as individuals it's impossible not to have some sort of... Not to let your knowledge of that person affect you. And I don't think that's necessarily a bad thing, as I've said. I think sometimes knowing the individual can be good because you know not to give them too hard a time when you're assessing or you know sometimes actually you need to be quite firm with them or you know they can do better.' (Trainer 12)

Thus, having prior knowledge of the trainee might potentially impact upon trainers' judgement, and some (as above) were prepared to admit to preconceived ideas about trainees, and the effect upon an assessment which might result. The effect of this might be mitigated by trainees having multiple assessments with many trainers who had differing perspectives upon their work, but there were individual interviewees from both groups who felt that the only way that they could see the assessments becoming less subject to this effect would be to use external assessors. An example of this from a trainee's viewpoint was;

'The marking has to be really fair because internal examiners they'll be biased and they've seen the candidates before so they might give me more marks even if they don't do well, whereas an external person who hasn't seen me, if he assesses me then I feel more confident that way.' (Trainee 6)

Though the trainee's comment is couched in terms of summative assessment, it does convey their feelings regarding trainers' prior knowledge of them. Although only this trainee advocated the use of external assessors, it does lend weight to some trainees' feeling that assessors' prior views about trainees might affect how they assessed them (for better or worse), and as the trainers quoted previously said, it seems very unlikely that a trainer's previous impression of a trainee would be completely forgotten during an assessment.

As suggested earlier, there were some trainees who clearly perceived advantages in engaging with multiple assessors, such as reducing the effect of prior impression, one suggesting;

'Increasing the number of assessors, I think, because I think just sticking to a cohort of consultants, three or four, in one particular attachment, and you have worked with them, so they know exactly how you work. And they might give you feedback based on that exercise and, to a certain extent, based on the fact that they have worked with you and they know you have done better on a particular day, whereas when you have a totally new cohort of consultants, who you haven't probably worked much with or interacted with, they might actually assess you in a more... on that day's performance.' (Trainee 8)

Another important issue is that although earlier quotations suggested that some trainees might look for assessors they perceived as generous, some trainees' apparent selectivity was also driven by trying to find trainers who would engage with the process. Some trainers were felt to be poorly engaged in the process, a trainee commenting;

'A few trainers aren't as interested in doing them so you end up having a little bit of a skewed assessment... But that just means that you end up picking and choosing a little bit who assesses you, which I guess isn't a fair representation, but the ones that have been done have been done well.' (Trainee 3)

A perceived lack of engagement on the part of some trainers was exacerbated for some trainees by practical difficulties in finding trainers who had time to undertake assessments during the working day, and this formed a major consideration for some trainees in deciding who to select;

'I think one of the main reasons, I would say, of seeking out is accessibility, I think. If I find a particular consultant is quite accessible, or is free at that point of time, I tend to approach him or her.' (Trainee 8)

The inhibitory affect of appearing too busy to undertake assessments was acknowledged by one of the trainers, who said;

'I think it's probably the people who make themselves available are probably more likely to be asked if they're shown to have the time and I suspect personally I probably feel I'm quite busy so probably don't open myself to being approached.' (Trainer 12)

Overall, in the day to day process of initiating WPBA, there was evidence that trainers may inhibit trainees' requests for assessment by both refusing to engage in the process and appearing too busy to fit them in. This demonstrates that the trainers' authority over trainees may be sufficient to prevent an assessment occurring. As a result, some of the features of some trainees' apparently selective approach to WPBA may be partly due to factors associated with trainers, as well as those perceived to be due to trainees themselves.

c. Summary of sections a and b

The preceding sections regarding interviewees' beliefs and actions prior to WPBA (to the point of initiating them) in routine practice demonstrate that there are varying views on what occurs during this phase, both between trainers and trainees, and within each group. Trainers appear to view trainees' motivation as largely subject to the need to meet numeric targets, and whilst other factors were also suggested by trainees, such as wishing to increase confidence, the influence of targets also featured prominently in their responses.

There was variable practice in initiating WPBA, due to instances of both trainees and trainers taking the lead, and there were examples of interviewees viewing both methods as satisfactory. Though trainees appeared comfortable in choosing cases for assessment, there were times when trainers wished to make that choice. There was more consensus regarding choosing cases for assessment prospectively, but even in this situation, there were a minority who would allow retrospectively chosen cases to be used.

There were clear instances of both trainers and trainees wishing to have prior knowledge of their assessee or assessor prior to an assessment, to the extent that one trainer would not assess a trainee of whom he did not have foreknowledge, whilst others advocated a more distant (or even external) relationship. The effect of being assessed by multiple trainers was suggested as a means of mitigating a single assessor's potentially subjective view.

These and other factors such as some trainers wishing to control the timing of the assessments, being poorly engaged or appearing too busy to assess trainees meant that the initiation of WPBA was subject to numerous factors and constraints beyond trainees simply asking to be assessed.

As a result there was not a clear unifying message regarding day to day practice during the lead into and initiation of WPBA. Both belief and action during this period were variable and subject to numerous influences. This meant that trainees did not always adopt a standard means of initiating assessments, but sometimes had to adapt to the constraints and preferences of trainers too. Differing users' expectations of the process could lead to variation in how trainers and trainees constructed assessment episodes, influencing their enactment, as discussed in the following section.

d. How assessments were enacted

I was interested in the interviewees' attitudes during the assessments, both with regard to the process itself and key aspects of the relationship between assessors and assessees. This section relates to interviewees' views on what actually happened during episodes of assessment, rather than how they prepared for them or events which occurred afterwards.

In this section, interviewees' attitudes were addressed by looking at how respondents thought trainees viewed the assessments' importance and how it affected trainees' behaviour during them. A view of the two parties' relationship during the assessment

was based upon how strongly trainers felt regarding whether or not trainees closely followed the assessor's own practice whilst undertaking WPBA.

Subsequently I looked at how both trainees and trainers might act during assessments, particularly how trainees might attempt to manipulate the process to their advantage and how some trainers might attempt to shorten or streamline assessments, adopting a peremptory approach to WPBA.

These areas could only represent sampling of the many facets of the enactment of WPBA, and were chosen to highlight attitudes informing both trainees' and trainers' actions during assessments, as I felt these issues represented the key factors influencing what occurred. Some (e.g. practising the trainer's way) were chosen because I had specifically asked about them during the interviews, having anticipated that the enquiry might yield useful information. Other areas, such as instances of perceived manipulation of the WPBA process by both trainers and trainees, arose because interviewees felt they were important, and chose to discuss them.

In terms of how trainees might view the assessments, Grant et al's (2007) study of the use of WPBA in foundation suggested confusion amongst trainees as to whether the assessments represented tests requiring a special effort or purely observation of routine practice. It is difficult to think that the presence of a trainer would have no effect upon a trainee's effort or performance during an assessment, even if it only caused mild anxiety, but despite this there were differences in how trainees felt they behaved during assessments.

A minority of the trainees interviewed were explicit about making a special effort for the assessments, one stating;

'I think my mind automatically goes into a test mode, and I presume when I'm doing a workplace assessment, I think I'm at my best, I can tell you that.... So you're all up in arms, you know, very conscious about you are being assessed.'
(Trainee 8)

This was the most overt expression of a special effort being made for assessments, and although others may have felt similarly, other trainees expressed contrary views, suggesting that the issue highlighted from foundation training, regarding whether or not a special effort was required, might also apply to radiology. An opposite view to thinking a special effort was required was expressed by this trainee, who said;

'I think it should be an everyday work. I think that's why these tools are there, because every day we should be good at it. It's not just for that particular case.'
(Trainee 6)

There thus seemed to be great variation in this aspect of how trainees undertook WPBA. Trainers made little mention of the 'special effort or not' issue, and although one acknowledged that a degree of performance was likely during an assessment, he related this to everyday work, and played down any adverse effect his presence might have upon the trainee's assessment (e.g. by making them nervous), saying;

'You are viewing their performance primarily but you are viewing something that is a reflection of their ability to deal with their everyday work because I am a less frightening person than a consultant surgeon.' (Trainer 11)

Whether they did it consciously or not, some trainers did adopt strategies which might negate trainees preparing for assessments by, for example, announcing an assessment at short notice, one describing her method of choosing a case thus;

'I say we decide that we're going to do it on a particular list, and then the list is booked as normal, and then almost at random, we just select a case.' (Trainer 3)

As well as illustrating the different ways trainers undertake WPBA, the quotation also demonstrates that for some trainers, opportunities to initiate assessments and lead or collaborate over case selection are things they already exploit, thus feeding back upon the varying day to day practices highlighted in the previous section.

In addition to enquiring how trainers might commence episodes of WPBA, I also wanted to investigate their views regarding trainees' practice during the assessments and how this might affect the trainer's actions. In order to do this, I asked them how they would feel if a trainee were to undertake a procedure their own way, rather than that of the trainer, during an assessment, assuming that both methods were safe. The majority of the trainers stated that they would be happy to watch a trainee use an alternative method, and two said that they had learnt new methods themselves through assessing trainees, one of them stating;

'I think that's part of the richness of learning, and indeed there are times when I will learn something from a trainee showing me another way of doing something.' (Trainer 5)

However, a minority of trainers took a different view, and felt that the trainee should follow their previously demonstrated method during an assessment, or face being marked down, as articulated by this trainer;

'When we're doing procedures I expect them to do it my way during that procedure, during my session. I say to them that there are many ways of doing intervention say, and they have to see them all and experience them all and then make their own mind up, but whilst they're in my session, and if I say to them, "I would like you to do it this way", that's what I expect.' (Trainer 6)

This quotation introduces the additional issue of authority relations between trainers and trainees, and was the most overt example of an assessor wanting it done 'their way'.

Another trainer articulated a similar view in slightly more moderate terms, saying;

'If I've told them already, I'd like you to do it this way, and they don't do that, well that shows I think, again maybe the reason most people do things in a certain way is because of experience and having seen how things can go wrong, and to just disregard that I think is worrying actually, that they can't follow instructions.' (Trainer 4)

Generally, trainers expressed a more relaxed view regarding 'their way', as shown by the following quotation;

'If that's a slightly different way to the way I would do it, so long as it's a logical way and is a reasonable way, I don't mind if it's not my way.' (Trainer 3)

As discussed in section b, but also impacting upon this issue, there was some concern expressed in the trainees' responses regarding the effect of previous trainer impression of them upon their assessments. One trainee expressed the view that trainers could be

influenced by their prior view or by the way they liked things done, whilst another admitted to adapting her method, depending on who was assessing her, saying;

'You learn as a trainee that different consultants like things being done in a different way, and often you adapt your approaches anyway in that way.'
(Trainee 3)

This could be seen as another manifestation of what might be termed gaming during WPBA, which was noted previously, and arises at several points within the analysis. However, it could also be interpreted as trainees being open to trying new methods, and both she and another trainee felt that learning from (and adapting to) different trainers' preferences led to improvement in the long term, as they incorporated varying consultant opinions into their own practice, a colleague saying;

'Everyone's got a slightly different way of doing things, a different take on things. And as a trainee, you can amalgamate those different ways of doing things to form your own.' (Trainee 5)

An important additional issue is that although the trainers I interviewed appeared to take their role as assessors seriously, some trainees felt that others adopted a more peremptory approach, one stating;

'I've generally had a positive experience but I've had a couple of people filling forms in not really giving me proper feedback and going through them quite quickly, and I've definitely seen other trainees having forms filled in with them not being there.' (Trainee 3)

This illustrates that perceived manipulation of, or opportunistic utilisation of WPBA to obtain a desired outcome (such as shortening the process) is not just a practice allegedly adopted by trainees, and that some trainers may engage in similar practices, as was initially discussed at the end of section b. Though it has been suggested earlier that some trainees may adopt a numbers-driven, perfunctory approach to the assessments, some trainers' behaviour may also contribute to this. An example of such an episode was provided by this trainee;

'I've reported a case and then I would just have to send them an electronic ticket. And then I got feedback, as in I got it into my system, but I didn't actually have to discuss anything with the case – they just looked at my report and decided that's good enough or not good enough.' (Trainee 2)

(An electronic ticket is a mechanism where a trainer is reminded to fill in an assessment form in a trainee's electronic portfolio following an assessment and may be done remotely.)

In some instances, when a perfunctory approach has been perceived to have been adopted, one might feel that there was an element of collusion between the two parties in order to bring an assessment episode to a speedy conclusion. An example of this was provided by a trainee who found she needed to reach a target number of assessments for her annual review (ARCP), shortly after the introduction of WPBA. She described what followed thus;

'We had this ARCP coming up and we'd been told we had to produce all this, it had to be from my side rather than coming from their side, so the initiative had to be from myself, and I had to tell them, "Oh can you just do this for these cases?" And they just simply did it for those cases.' (Trainee 6)

This demonstrates that (in this instance) a trainer was happy to help a trainee build WPBA numbers quickly, without either party appearing to fully engage in the assessment process. It may have resulted in the construction of a process which suited them both, but might not have best promoted trainee development in terms of how the WPBA episodes were enacted.

Factors such as a perfunctory approach, refusal to engage in WPBA, and inhibitory behaviour on the part of some trainers demonstrate that the negative perceptions of some trainers regarding trainees' engagement with WPBA are in some instances mirrored by trainer behaviours which encourage or influence trainees to adopt the approach they do. Thus members of both groups felt that there were examples (in the opposite group) of poor engagement or commitment to WPBA, as well as areas of good practice.

In summary, the assessments were not enacted uniformly. The trainees' attitudes ranged from viewing WPBA as an evaluation of routine practice to making a special effort for the assessments, whilst those trainers who wanted trainees to perform their way were balanced by larger numbers of colleagues who were comfortable with trainees adopting other methods.

A trainee mentioned adapting her method depending on who was assessing her, and although this might be seen as an example of gaming, it could also be seen as using the input of multiple assessors to build future practice. This was followed by examples of trainers themselves taking an opportunistic or even perfunctory approach to WPBA, and again serves to highlight the heterogeneity of day to day practice, although such

differences between both these, and other practices, might be viewed as variations upon a central theme, rather than a completely different method.

e. How feedback was delivered

As noted in section a, there were varying views expressed regarding whether or not trainees reflected on WPBA following their assessments. An important aspect of practice which contributes to this is the delivery of feedback and its purpose as described by Hattie and Timperley (2007) is reducing discrepancy between current practices or understandings and those which are desired. The same authors analysed numerous meta-analyses of feedback, and concluded that its highest effect occurred when students were given information regarding how they had performed a task, with additional guidance on how to perform it more effectively.

Feedback's greatest effect occurring with improving the performance of tasks suggests that such a technique would suit a specialty such as radiology, where much practical procedural work (i.e. specifically task-based) is undertaken. To best achieve such a good effect, guidelines for the ideal delivery of feedback suggest commencing on a positive note, focusing on behaviour which may be changed, including factual objective comments, making constructive suggestions, avoiding value judgements, and developing an action plan (Fullerton, 2003).

The previous paragraph concerns how feedback is delivery is moderated, and other important attributes of this include its timing (immediate or delayed), specificity and whether it is delivered face to face or remotely. Norcini and Burch (2007) describe how

feedback's efficacy is improved when it is both timely and specific, and both attributes are addressed by interviewees in this section.

Overall, this section concentrates upon feedback as part of the assessments, addressing interviewees' opinions regarding its importance within day to day WPBA, their views regarding its delivery, its place within a continuous educational process, and its perceived effect upon trainees' future learning in an attempt to gauge their feelings regarding the practical impact of WPBA. Finally, some interviewees mentioned instances where trainers were thought to have given over generous feedback and these are exemplified and discussed.

The trainees' views reflected their opinions presented in chapter five, in which they regarded WPBA as formative with feedback delivery an integral feature of the process. Several trainees expressed the view that the assessments were a stimulus to both reflection and future learning;

'I did the procedure under the supervision of the consultant who was assessing me, there would be learning points and learning triggers that could come out of that assessment.' (Trainee 1)

and

'I think it has improved the way I'm learning and the way I perform in the future. It's made me far more inspired in a way to go and actually read up on things and try and improve things, whereas prior to doing these I used to think things were going along okay, something went badly we'd maybe have a little bit of a think about it and move on.' (Trainee 3)

Both of these quotations came from highly motivated trainees who perceived that the day to day use of WPBA would stimulate further learning in the future. However, two

other trainees, who could also be described as motivated and successful, felt that the introduction of WPBA had had little effect upon their future learning. The two trainees who voiced these negative opinions regarding the assessments did value the provision of feedback (which they were already getting) but felt that the new assessment process had not improved that which they were already receiving, one stating;

'But, on the whole, the vast majority, on a day-to-day basis, they're constantly giving feedback already without the workplace-based assessments. And, in my experience, I don't get any more feedback from them than I do in normal routine practice.' (Trainee 4)

By contrast, the two trainees quoted at the start of this section appeared far more enthusiastic about the utility of the feedback they received as part of WPBA, and there were those who felt such useful feedback would not have been as easily available prior to the introduction of the scheme. In some trainees' view it had led to a particular improvement in the specificity of their learning, as shown in the following examples;

'I find it very useful because, as I say, I've had feedback that I don't think I otherwise would have got on specifics. So it's been very useful from that point of view in terms of learning.' (Trainee 7)

and

'We did our first workplace based assessment, the DOPS, it was pretty early on after I'd started doing some biopsy and breast work, and we raised a couple of points not necessarily around about problems that were seen but maybe problems that might arise in the future, and those definitely were reflected in how I went about doing biopsies in the future.' (Trainee 1)

A quotation from a trainee in the previous chapter suggested that WPBA helped the delivery of feedback by formalising the assessment process and mandating the assessor to deliver it (as per the College's guidelines), but the latter exchanges suggest that the quality, and importantly, the specificity of feedback may have been improved by the scheme. In addition, the trainee quoted latterly on the previous page was specifically shown how to improve his future practice.

Another trainee was also enthusiastic because WPBA provided him with written feedback on his performance, which he regarded as an improvement upon the oral feedback he'd received before, saying;

'When it's written and you read it... Because, sometimes, I get told, "Okay. This is where it went wrong" during the reporting session – you just pass through it, just in another information that's part of your daily learning. But when it's written there and you can reflect on it, it's better.' (Trainee 2)

Although this trainee was the only one to comment upon the value of the provision of written feedback as part of WPBA, other trainees were enthusiastic about the opportunity for face-to-face discussion which they felt was provided by the scheme, as demonstrated by the following quotations;

'The way we've done the workplace based assessment has meant that actually it's been more of a discussion about how it's gone, what we thought, why it was this and not this. You know, it's been a very much a discussion based, feedback centred event.' (Trainee 1)

and

'For me, the feedback actually, from what I receive from the consultant is the most valuable aspect of it. So I don't think it's a case of rushing it, but I don't want to sit next to somebody who's spending fifty minutes just staring at a computer screen, and not talking to me. The more useful bit, I find, is the bit

where they're talking to me, and it's recorded of a sort on the screen, but the conversation is the more useful bit for me.' (Trainee 5)

These quotations demonstrate the trainees' preference for discussion with the assessor following WPBA and the importance of such discussion as part of the delivery of feedback is reflected in Hauer and Kogan's (2012) view that trainee feedback must be included in the exchanges surrounding shared work, rather than delivered as a written evaluation after the event.

The converse of the perceived utility of the feedback delivery described above is the delivery of feedback remotely and/or after a prolonged time period. Although the precise timing of feedback delivery is not addressed in the WPBA guidance, trainees appeared to greatly value immediate feedback and the chance to discuss their work, and when this did not occur some trainees voiced their disappointment;

Q: 'Why is that a bad idea?'

A: 'Because it becomes like ticking boxes. It becomes like ticking boxes. How could you say I'm able to comment on something, because we didn't have the conversation. I think sending electronic tickets, for these, for direct observation skills or for that, because, mainly, the main reason is you don't get the feedback.' (Trainee 2)

and

'I just felt what's the point doing a form? So you just do it and you send it and it's back, so it's just a bit of paperwork, that's what I felt when I didn't get any feedback from some of them.' (Trainee 6)

Negative comments regarding the receipt of delayed or remote feedback were also made by other trainees, and it is clear that the quality of feedback and how and when it was

delivered were important factors in many trainees' judgement of the day to day utility of WPBA.

The impact of feedback delivery could extend beyond the technical aspects of professional development. For one trainee it impacted upon their sense of being a full member the imaging team, as demonstrated thus;

'I feel that I've had more interactions with some of the consultants, we've always got on quite well as a team, but actually it does mean that at the end of a day or the end of a session at lunchtime you do end up sitting down together with a cup of tea and going through the form, so it's increased interactions with some of the senior members of the department I think, which has been quite positive really and it makes you feel a bit more of a member of a team.' (Trainee 3)

Although this was an isolated (and unexpected) comment, it demonstrates another perceived benefit of the face-to-face delivery of feedback, where the trainee's relationship with other members of the team has been enhanced in a broader context.

Generally trainers concurred with trainees' views regarding feedback, particularly regarding its usefulness and the desirability of delivering it promptly and face-to-face, as evidenced by the following quotations from two trainers;

'I think, particularly because the immediacy of doing the assessment after the trainee has actually performed a task, they do get rapid feedback and the feedback itself is, I think, quite comprehensive.' (Trainer 1)

and

'I think it's important to give it straight after, as soon as possible, but I think it has to be the same day. I think it's... we are so busy it'll be very difficult to do it the next day or even a week later. It should be ideally straight after the procedure, take five minutes to sit back and discuss it out.' (Trainer 8)

Although no trainers I interviewed advocated different timescales or remote (electronic) methods for feedback delivery, it was clear from trainees quoted earlier that delayed and distant delivery did occur, and were not appreciated by those who chose to comment. With regard to this issue, it is possible that the trainer sample comprised individuals motivated to deliver feedback immediately and face to face, and had they used other methods, they may have chosen not to discuss them.

In addition to the issues discussed above, and as another aspect of feedback delivery (perhaps reflecting some trainees' desire to build their confidence from the assessments), some trainers were keen to stress positive aspects of a trainee's performance during feedback delivery. An example of this is as follows;

'I believe very strongly in positive reinforcement in improving confidence and taking away any element of anxiety and fear that they have about their daily activities.' (Trainer 11)

Thus in some instances, it appeared that trainers and trainees might work together to deliver and receive positive feedback in order to try and boost trainees' confidence. At no point did a trainer say that they gave positive feedback following a poor assessment, but examples of scoring perceived as over generous are discussed later in this section and the next chapter contains instances of trainers committing less critical comments to the record than were made orally at the time of WPBA.

An additional facet of feedback delivery commented upon by trainees was a desire to receive specific guidance. Some trainers were also keen to emphasise the importance of

delivering a specific and clear message as a means of improving the utility of feedback, as illustrated below;

'For most trainees I think it's just a nice way of being able to say, well we did that biopsy today or we looked at these chest CTs today, and you kept forgetting to look at the old CT, you kept forgetting to window it, and just to reinforce I'd like you to do that next week and have it documented.' (Trainer 4)

and

'I've had a very competent Nepalese gentleman who I've been responsible for various aspects of training. I've had a very competent Greek Cypriot, but their accents are strong, they are dealing with an elderly population often Caucasian and English, and one thing that I will often... a common mistake is to stand in the room rather than sit, and there are various aspects I repetitively teach, and although all these are very good trainees they need to improve their communication skills and I've sat down and talked to them very openly about this, speaking slowly, how to minimise the accent, making sure their body language is correct and sitting down on a chair instead of standing over somebody dying of cancer.' (Trainer 7)

The importance of delivering a clear message when giving feedback was emphasised by a trainer, who recalled episodes when she was a trainee, demonstrating how unclear messages reduced the perceived utility of assessments;

'I could have two consultants giving me two different views, which used to be a few ticked boxes but a lot more of free text; and they never married up, the two things, I would spend a lot of time wondering what each person is trying to get at.'

She then added;

'I think that's one thing which I, as a trainee, found it lacking in the system was: people would tell me what was wrong and where I was wrong, but they never actually told me how I could improve' (Trainer 6)

These negative opinions serve to demonstrate how poorly delivered feedback may impact very adversely on assessments. Aside from any formal training they may have received, some trainers also found the assessment scheme itself made them consider how their feedback might be improved, one trainer saying;

'they probably give you an opportunity to focus a bit on what... exactly what it is trainees are being expected to learn. And maybe some of the things outside of the... I mean, I'm talking particularly about procedural assessments, so actually spending some time listening to how they interact with the patient and so on, as opposed to just being interested in what the images look like in the end and what their final interpretation might be.' (Trainer 3)

Amplifying his view expressed in section a, one trainer discussed how he felt that useful feedback might have its utility increased even more, by trainees applying it to multiple cases in the future, thus multiplying the effects of assessments, saying;

'If they can translate some of the specific learning activities or the main gist of the specific learning activity, and then translate it into their general practice and their everyday work and so on, it obviously becomes quite an integral part of their learning process. And I think that's the way they should be viewing it. And if that's the way you look at it, then even if we did, say, just five Mini Interpretation Exercises on bone scans in the attachment, they'll probably multiply that learning process several times.' (Trainer 2)

This quotation serves to illustrate the view, expressed by some trainers, that although WPBA consists of single episodes with (often) different assessors, they should be integrated into a continual learning process in order to maximise their utility. The converse of this is demonstrated by the following quotation, in which a trainer expresses a view that some trainees see single episodes of WPBA as end points, and do not integrate them into a continuous learning process;

'Yes, because say I have done a trainee's ability to report acute CT abdomen, or acute CT of the aorta, and I've said I've given them a pass equivalent, if you're going to call it as a pass or fail otherwise – because I don't believe it as a pass or fail. Say if I said they're on par with their peers and they require only minimal indirect supervision: that may suddenly sink into them and say "Right, now I know how to do a CT aorta and it's on paper, it's on print. It's there with the deanery now, so nobody can change that." So they may not put in more effort to learn more' (Trainer 6)

Aside from these examples, there were few negative comments from trainers regarding the contribution of feedback to the WPBA process, though in common with the views expressed by two trainees earlier in this section, some felt it did not add much to what had gone before, two trainers stating;

'It's not a new thing to feed back to the trainees how they're doing. Particularly for practical procedures, we'll have a discussion at the end of each session or each case to say how did that go, where did things go well or badly? So we were doing that before, this is just, I don't think it's actually changed that, I think it's just structured it a bit more, and documented it a bit more.' (Trainer 8)

and

'I think, in terms of measurable outcomes, it won't improve overall quality very greatly, because I think, even trainees that are borderline will still be borderline and they'll still come through at the end.' (Trainer 1)

This represents an important illustration of the fact that both the delivery of feedback, and its effect upon learning in radiology, are not new issues, and some interviewees saw WPBA as just a means of structuring and formalising the previous process. The second quotation suggests that the interviewee felt that the assessment scheme would have a minimal effect upon training, if any impact at all. This might be because the respondent

felt that training pre WPBA was satisfactory, and although it was not improved by the assessments' introduction, it did not deteriorate afterwards.

Although this is likely to represent a wider issue than WPBA and delivery of feedback as part of the process, another possible problem mentioned by a minority of interviewees was their opinion that trainers might deliver feedback which was over-generous, with a tendency to stray towards good scores when performance was satisfactory, a trainee commenting that;

'I think trainers struggle as well at the moment to put a tick box in the middle which is "meets expectation". They want to stray towards the right hand side which is, you know, "above expectation", because that's where the good tick would usually have been in the normal scales that we've used previously.'
(Trainee 1)

There were also trainers who saw this as an issue, one saying;

'I think maybe workplace based assessment can be over-favourable to the trainee. I think most assessors struggle with scoring people down on any assessment, so I think that a lot of trainees will get by with satisfactory, across all of their scores, and in fact what that will mean, having only satisfactory as your score means that actually there are concerns.' (Trainer 4)

This may be due to assessors wishing to appear supportive, allied to feeling uncomfortable with delivering a negative or just satisfactory message. Such discomfort is illustrated by the following statement from a trainer;

'I think it's always difficult breaking some kind of bad news to anybody, isn't it? So I think when there are issues and you're almost being confronted with having to discuss through them with the trainee... And so in fact it's probably the most helpful to the trainee way of doing things. It's probably harder for the trainer because you're confronted with having to do it and having to be honest.'
(Trainer 10)

Difficulty in relaying negative feedback resonates with the desire expressed earlier by some trainers to emphasise what went well, and for trainees to have their confidence increased by undertaking WPBA, but may come at the cost of feedback being generous rather than strictly accurate. Interestingly, no interviewee suggested that feedback had been delivered which they judged harsh or over critical, perhaps reflecting an ethos of encouraging and supporting trainees through WPBA.

Although positive reinforcement may help trainees develop their skills, the reasons for trainers' delivery of generous feedback were investigated by Kogan et al (2012), who suggested that because of trainers' own psychosocial need not to appear unkind, trainers' perception that trainees' self esteem should be preserved, and their concern regarding the maintenance of the trainer-trainee relationship meant that assessors sometimes struggled to give negative feedback. This view was supported by Crossley et al (2011) who suggested that trainers were generally indiscriminate, and rated the majority of trainees very positively.

In terms of addressing this issue, Kogan et al (2012) suggested that it might be helped by reminding faculty that their feedback followed assessment for (rather than of) learning and ensuring that feedback is delivered within an agreed framework in the context of expected milestones in training. To some extent this is addressed in the Royal College of Radiologists' scheme which asks findings to be related to a trainee's expected stage of training.

Overall, as with other stages of the day to day assessment process, it is not possible to formulate a consistent message with regard to what occurred during the delivery of

feedback. Interviewees were able to express preferences for feedback that was clear, specific, delivered face to face and timely, representing elements of established good practice. There were certainly instances of such feedback being delivered as part of day to day assessment, but also examples which were judged by trainees to demonstrate lack of clarity, remoteness and delayed delivery.

Some trainees clearly valued the feedback they received and felt it aided their future learning, but there were both trainers and trainees who felt that this process did not represent an improvement upon the informal delivery of feedback which occurred prior to the introduction of WPBA. In addition there were trainers who questioned how much trainees reflected upon their feedback and whether they integrated individual assessments into a coherent learning pathway.

Another trainer expressed the pessimistic view that WPBA and the delivery of feedback would not improve the outcome of training, but did not say why they held this view. The measurement of training outcomes would be very difficult, aside from analysing short term data such as exam results or compiling users' views regarding the effect of the assessments. A lesser issue raised by a minority of interviewees was a tendency for some feedback to be too generous.

It is clear that the central guidance for the conduct of WPBA regards the delivery of feedback as an integral part of a process of formative assessment, and the interviewees' responses aligned with this, with none questioning either the desirability or need for feedback as part of the assessments. Although there may have been inconsistencies in how the trainees received feedback, reported instances of either no or very delayed

delivery were uncommon, suggesting that users accepted the integral nature of feedback and almost always included it.

7: Theme 3:

How users influence the realisation of WPBA

This chapter addresses how the realisation of WPBA is consciously influenced or adapted by the participants. After an initial review of the data, I decided to include the following material in the current chapter:

- summative usage of the assessments
- different uses of WPBA by trainers depending upon the trainees being assessed
- amendments to how the assessments are recorded
- assessing trainees without their prior knowledge

It is important to note that there are potential areas of overlap between the previous chapter regarding the day to day realisation of WPBA and the current one about users' influence upon the process. The very act of trainer and trainee coming together to construct an assessment inevitably means that the two parties will influence what results.

Clearly it is difficult to separate influences users bring to bear upon the process from how they construct WPBA during its day to day enactment, and a judgement was made upon the separation of material between the current chapter and the last. This chapter includes examples of where I felt users had deliberately adapted the process (particularly

on the part of trainers) rather than what occurred during their day to day enactment of the assessments.

At the start of chapters five and six, particularly the latter, short summaries of central policy outlining the formative ethos of WPBA were included, and it is important to refer back to these when trying to assess how the conduct of the assessments was influenced by users. Guidance and score sheets for undertaking WPBA were issued by The RCR (2010) and these are included as appendices g - j. In conjunction with the college curriculum, they provide fairly detailed guidance regarding the conduct of assessments, and thus might be expected to guide how WPBA is enacted in the workplace.

a. Issues of adapting the assessments for summative use

It is clear from the previous chapters that the tension between summative assessment and the intended formative use of WPBA is a central issue. In this section this tension is analysed from the broadly contrasting perspectives of trainers and trainees. The mechanisms by which the tension may be resolved are also examined. An example of the latter discussed in this chapter is the possible use of a collection of WPBAs to determine trainees' engagement with the process, inform summative judgements, and help determine if trainees may progress in training.

Chapter five demonstrated that a greater proportion of trainees than trainers felt that WPBA was for formative use only. This view was conveyed by a trainee who said;

'I think it just tells you at what level you are, and it highlights your weaknesses. And so then, you develop on that. You reflect and develop on that. So I think it's a completely developmental process. I don't think it's a final note.' (Trainee 8)

A minority of trainers were similarly emphatic about the formative nature of the assessments, one saying;

'My view's that they are misrepresented currently, they're used as summative assessments, which they are not, and not designed to be, instead of being used as formative developmental tools.' (Trainer 5)

Whilst another felt that WPBA was too subjective for summative use, saying;

'I think that the problem with workplace assessment is that so much of it is subjective, and the move really has been much more to objective, reproducible, egalitarian assessment.' (Trainer 11)

The second trainer perceived a weakness of the assessments which he felt made them unsuitable for summative use (for which they were not originally designed). His quotation raises issues with respect to WPBA's perceived reliability, and this is likely to form one of the factors behind this trainer's view. However, other trainers were prepared to use them both formatively and summatively in certain circumstances, one stating;

'I think, if someone's coming towards the end of their training period and they were doing a variety of examinations and you think they were doing them as well as you would hope yourself, then I think they can be used as a final assessment as well.' (Trainer 12)

There are already alternative methods for summative assessment (such as external examinations, see appendix f) in place, although it might be argued that these do not directly assess workplace performance. There was general agreement between both

groups that assessment of a single episode could not be used for summative purposes, as illustrated by the following quotations from a trainee (first) and a trainer (second);

'I think a one-off workplace assessment should not be used as a final, that's my personal opinion. I think there should be a number of workplace assessments over a period of time that should be used to assess that this person is competent.' (Trainee 8)

and

'I wouldn't want to pass or fail somebody on a single workplace based assessment. I think there would have to be a series of them.' (Trainer 3)

Amongst those who thought summative assessment possible at all, there was a general feeling expressed by interviewees that a summative judgment could only be based upon a collection of assessments, reflecting (but not necessarily directly based upon) the opinion expressed in the report from the RCR trial (Royal College of Radiologists, 2010) quoted at the start of chapter five. Both the interviewees' opinions and that of the trial seemed to be based upon the view that multiple assessors, judgements and episodes could give a more reliable assessment of competence than a single episode, and a trainer who was involved in the development of WPBA conceded that;

'You can use multiple Workplace Based Assessments on multiple topics by many observers collectively to make a summative judgment. You can also use evidence of engagement and quality of engagement with the process to give an indication of a trainee's participation in the educational contract.' (Trainer 5)

There was more limited support for use of multiple episodes for summative assessment amongst the trainees, with greater support for the overall consultant comment passed at

the end of a period of training (discussed in section b of the following chapter). An example of the limited support for WPBA in the role of summative assessment was expressed by a trainee who felt (as did many others) that the assessments were primarily formative in nature, but then added;

'I think it can be used finally as part of an overall assessment. And I think it's already being used for that, to be honest. I think if your Educational Supervisor looks at all of your workplace based assessments and has to write a report on you based partially on those, I think it already is being used as a final outcome.' (Trainee 7)

However, the majority of the trainees emphasised disadvantages in using WPBA for summative purposes, such as perceived variability in the standards different assessors might expect and the additional pressure of performing on the day, although this represents a different situation (a single episode) to that envisaged in the previous quotation, where a supervisor might look at a series of assessment results. The perceived additional pressure of performing well on the day is exemplified by the quotation which follows from this trainee, though it could be applied to any one-off summative assessment, and not just WPBA;

'I think there'll be an extra pressure really. I mean all these exams, first of all our performance really will go down when someone is assessing you. I just feel that if they say it's a pass or fail then there'll be an extra pressure on us to perform really, really well, I might be really good at it but I might ruin the whole thing just because I'm being assessed.' (Trainee 6)

Within the minority of trainees who said WPBA could be used summatively if the judgement was based upon multiple episodes, some added the further caveat that the

assessments should be undertaken over a long time period, the latter being highlighted by the following quotation;

'You'd be showing that you've done a number of them so not just two assessments a day before something's due in, that you've been doing them over a period of time over the month, so with a number of assessments being performed each month, and showing that you've gained competency over the years.'
(Trainee 3)

Even though some trainers appeared keener on the assessments being used summatively than the trainees, some who thought that the assessments could potentially be used for final sign-off also felt the need to qualify their answers, one saying;

'It would need to be made very clear to the trainee and the trainer exactly what was required in terms of the number and the case mix within that number that had to be achieved. In that case the numbers would need to be substantially greater in my view and the time component would need to be built in to the work pattern of the trainee and the trainer.' (Trainer 11)

Current arrangements for trainees being certified as having completed their training are based upon a system of annual reviews to determine satisfactory progress and passing external examinations. Interviewees who discussed the possible use of WPBA in this situation mentioned numerous qualifying factors including the need for standardised cases, multiple assessors and a wide breadth of assessment scenarios, almost taking the assessments into the realm of an OSCE (objective structured clinical examination), where a structured marking system is used to assess trainees undertaking standardised tasks during a circuit of assessment stations (Pell et al, 2008). Such arrangements are

used for summative assessment in medicine, although not final certification, and to this end, one trainer suggested that;

'You'd probably need to bank cases to do that rather than just use a general work list which might not cover the full scope of cases that you would want to know that someone's competent at.' (Trainer 9)

It could be argued that the above represents such a departure from the formative ethos of WPBA that it represents an entirely different form of assessment. One trainer seemed prepared to extend the potential summative role of WPBA even further, suggesting that a portfolio of assessments might be used for final sign off of trainees, or as supporting evidence when they applied for consultant posts, stating;

'I think that's part of your evidence and it will become, you know, accepted as part of your eportfolio and that you may use elements of that in job interviews and in shortlisting.' (Trainer 5)

In certain circumstances there were also some trainers who would use single assessments in a summative fashion depending on whom they were assessing. Examples of this were demonstrated by small numbers of trainers who felt the purpose (and their usage) of WPBA might alter during a trainee's attachment or according to a trainee's seniority. An example of the latter was expressed by this trainer, who said;

'You see, again for the procedure I think is to have a competency to actually say if they can perform it on their own, at least for the older students. I think for the younger ones who haven't done many perhaps we just want to assess at what stage they are and see what needs they have.' (Trainer 8)

This is an important point, as it represents an example of the same assessment being used differently when assessing trainees of varying levels, and both this and the instance which follows are important exemplars of how WPBA may be influenced by trainers when it is realised in practice. Another view expressed by a minority of trainers was that their assessments would become more summative in nature as a trainee neared the end of an attachment, a trainer stating;

‘The first time I do a workplace-based assessment with them, I tell them, “Well, at this stage, I’m not really that concerned as to whether you’re getting it right or wrong. I want to see you demonstrating a learning process and a process of development during the attachment, so I’m hoping that if I showed you a case of a similar level of difficulty at the end of the attachment, you’d be able to bat it off.”’ (Trainer 2)

These represent comparable situations, as in both, trainers appeared to be influencing WPBA to be used as a means of signing-off trainees, either as trainees neared the end of their training or at the end of an attachment, using the assessments summatively rather than formatively.

In contrast to these trainers influencing the conduct of the assessments so they were used summatively for more senior trainees, one trainer was prepared to use the assessments summatively for junior trainees, but not those who were more senior (she did not state any defined point of transition). This was due to the difficulty she perceived in using the assessments to evaluate the nuances and subtleties of advanced work, and the trainer concerned explained her approach as follows;

‘I expect junior trainees to know the basic, the core, aspects of the subject. Whereas senior trainees I expect them to know a more detailed aspect of exactly

the same subject... That's when you look at the training more holistically, from all angles, and it's not just as a trainee who's come in "Hey well you've done this, and I've ticked your six boxes and finished." As I said, the procedures are an art and now I like to see the artist.' (Trainer 6)

This assessor preferred to comment freehand when assessing performance she saw as subtle and nuanced, rather than be constrained by the format of WPBA when assessing senior trainees. Although the assessments' score sheet does contain space for free text, much of it is structured as a check list where trainers rate the trainee's performance against that expected for their stage of training. Another trainer concurred with the view that the score sheet imposed constraints upon what could be recorded, saying;

'You can comment on all of those elements, but I don't think you can within the constraint... or I don't you can usefully within the constraints of the current form. Too many of the elements are tick box and too little is feedback.' (Trainer 5)

Although these comments might be seen as relevant to the utility of WPBA, they are also important with respect to trainers' influence upon the process, as they sought to escape the constraints of (what they saw as) ticking boxes. The examples demonstrate that some trainers see the means of documenting assessments as restrictive under some circumstances and that they influence the recording of WPBA in an attempt to deliver a more subtle or nuanced message.

This suggested a degree of frustration on the part of some trainers in delivering their opinion of trainees' performance within (in part) a closed scoring system and the issue is discussed further in section a of the next chapter, which concerns the utility of WPBA in

assessing complex work. In addition, issues regarding trainers either not recording assessments or delivering different verbal and written messages are discussed in the next section of this chapter.

b. Recording WPBA

Aside from some trainers' willingness to influence the assessments so they might be used summatively, in some circumstances, smaller numbers of trainers were prepared to amend the process in other ways. There were instances quoted where trainers would not document assessments, or others where what they said and subsequently wrote differed. The importance of this is that assessors' written comments form part of a trainee's permanent training record, and those who use such records to assess trainees' progress need to have confidence in their veracity.

One trainer suggested only recording favourable outcomes, with the purpose of avoiding damaging trainees' confidence. She compared her previous practice of informally assessing trainees at the start of their attachment with that applied currently (as part of WPBA), saying;

'I think giving them in their assessment not as good marks as they want is quite frightening for them, even though we've probably done it for years but we haven't written down at the beginning of your three-month block, you know, you weren't as good, or, really, you can't do fluoroscopy. But that wasn't written down whereas now we're meant to write down they're not good at the beginning so we can show improvement at the end' (Trainer 8)

She suggested that she'd prefer a less formalised system of assessment where trainers did not dwell upon suboptimal episodes en route to developing competence, but instead recorded trainees' success when skills were achieved. Thus she would still use WPBA formatively and deliver constructive feedback, but would not record a trainee's early steps to (hopefully) achieving competence. Another trainer adopted a similar technique by stating that although he might tell a trainee an assessment had gone badly, his written comments in the trainee's record would be more moderate, stating;

'If a trainee's done a case very badly, I will say, you know, "Could do better. This is what he's done well; this is what he hasn't done very well." So I may verbally say to them, "Well, that was actually a disaster." But I won't write down on paper that that was a disaster. I'll write down what they did well and what they didn't do so well, and how they can improve. So yes, I think there is a... I don't think it's a complete disconnect between what's being said verbally and what's being written down. But I think it usually does... a written down version does tend to be a slightly more mellow version of, perhaps, what was discussed.' (Trainer 2)

This comment suggests disinclination to permanently record episodes viewed as particularly suboptimal which may mean that trainees (and those subsequently involved in their training) have an inaccurate record of previous WPBA. The trainer involved did not explain why he would do this, but it is possibly due to his not wishing to appear unkind or undermine trainees' confidence by leaving a permanent unmoderated record of suboptimal performance. This might be seen in a similar way to not recording poor assessments or giving generous verbal feedback, as discussed in the previous chapter.

A more nuanced example of trainers not recording their true feelings in a trainee's record is demonstrated in the following exchange;

Q: 'Do trainees then not know what you really think?'

A: '... I think most people have good insight and actually are more critical of themselves than other people. I think they also understand that if they just get satisfactory then actually that isn't satisfactory, that they should actually be doing better than that.' (Trainer 4)

In this example the trainer appears to expect that trainees will understand that a 'satisfactory' result in their portfolio really means 'unsatisfactory' and that a trainee's own powers of self criticism will ensure that they will realise if their performance is substandard.

Assuming trainees will derive a different message from what is written, differing verbal and written feedback, and not recording poor assessments all represent issues of communication by which trainers may influence what is recorded, and if records are inaccurate or incomplete, it may lead to issues concerning the reliability of the whole process.

c. Assessment of trainees without their knowledge

Another way in which some trainers influenced the realisation of WPBA was to tell trainees that they were being assessed after the event. This approach is described by the one of the trainers using it, who said;

'I had the trainee for a whole list and said... they've had different cases and cases that don't happen very often like micturating cystograms, we haven't really said before, you know, that should be an assessment, but because he did well we said, "Would you like it to be an assessment?" Perhaps that's not the right way and you go like, should it be in secret, but that's how we used to do it many years before, you just watch your trainee and see how they're getting on.' (Trainer 8)

This could be seen as a trainer taking advantage of a spontaneous training opportunity which might have not been identified in advance. Not all opportunities for good training may be pre planned, but in terms of formally recorded WPBA both trainers and trainees might have an issue with this approach, particularly regarding the fairness of engaging in assessment ‘in secret’, as was acknowledged by the interviewee herself. She partly addressed this later in the interview, saying she would only do this if things had gone well, as shown in the following quotation;

‘I don’t think I have a problem if they do well, but they may have a problem picking the ones that they did badly, but again I feel like not all bad experiences should be recorded. I think it’s good enough for them to realise it’s bad and they haven’t got a good record so we don’t see in their folder they can do it.’
(Trainer 8)

As noted in this quotation and the previous section, this trainer took the view that not all suboptimal episodes should be recorded, but this would mean that only the retrospectively chosen episodes which went well would be the subject of recorded WPBA, and that the influence of this upon the assessment process might be that important opportunities for recording constructive criticism might be missed. As was also stated in the previous section, this is likely to impact upon the reliability of the process.

Another trainer adopted a similar strategy, justifying her view in the following exchange;

A: ‘Yes, because tomorrow when the trainee goes in and does the cases on their own, every case will be different and they have to independently

make all the decisions which are needed through the whole of the procedure.'

Q: 'When will they know it's an assessment?'

A: 'Afterwards; after the case has finished irrespective of the outcome.'
(Trainer 6)

She felt that such an approach ensured trainees were assessed in real-life situations, and enabled her to view their actions and responses in a more naturalistic setting, without any bias in selection of the case.

Very few of the trainees addressed this aspect of influencing WPBA during their interviews, although one of the senior trainees passed the following view;

'Sorry, getting assessed when you don't know you're being assessed, I don't think it would hurt. I don't actually, I think it would be fine. In a way we're getting assessed like that all the time. I mean, trainers are working with you and they're always... Whilst you are with them you're their responsibility. And so they're always looking at you and assessing you, I think.' (Trainee 7)

He appeared to take the view that unannounced assessment of trainees was occurring day to day almost as a matter of course, and seemed relaxed about this principle being applied to WPBA, although his perspective was that of a successful senior trainee who has since become a consultant. However, if episodes of WPBA are to be recorded and used as the basis for formal feedback, it is likely that other trainees might wish to know they are being assessed.

d. Summary

Although this chapter and that which preceded it have provided examples of how users may influence the WPBA process, it is important to note that these represent the more overt and consciously applied influences. In addition there are likely to be numerous subconscious influences upon the process (e.g. trainer's mood, time pressures etc) which due to their subtle nature are unlikely to be reported.

The topics discussed under the theme of how users influence WPBA's enactment inevitably show some overlap with those raised in the previous chapter regarding the day to day realisation of the assessments. The subjects derived from both chapters may be subdivided into those predominantly influenced by trainees, those associated with both individuals (assessor and assessee) undertaking an assessment and those influences exercised by trainers.

The main influence exercised by trainees upon the process appears to be at the start of the assessment, when they hold the initiative in asking for a WPBA to be arranged, selecting the case and assessor, and agreeing when the episode will occur. Although a trainer may not agree to all a trainee asks, or prefer to take the initiative themselves, at this point the trainee generally has the opportunity to greatly influence what will follow.

Thus early influence upon the process mainly relates to who initiates the case (case selection etc may then follow and fall to the initiator too) and also, importantly and inevitably, the trainee and trainer's prior knowledge of each other. Although initiation of the episodes is predominantly the duty of the trainee, there were interviewees who

advocated either joint or trainer led initiation of WPBA, and this topic is discussed further in chapter eight. The effects upon the assessments of the authority relationship between trainer and trainee, and how power may shift between them during an episode of WPBA are further addressed in the discussion.

Conscious influences brought to bear upon the process by trainers may be divided into three categories, these being summative usage of WPBA, issues concerning the recording of the assessments and assessing trainees without forewarning them. Both summative usage of WPBA and the recording of the assessments were subdivided.

The former was divided into trainers' and trainees' views regarding the general circumstances when it might occur, the identification of poor performers (from the previous chapter), and specific instances when some trainers might use a single assessment summatively. The latter was divided into leaving some assessments unrecorded, differences between written and verbal feedback, and recording comments which the trainee was supposed to interpret differently.

Aside from the frank adaptations described above, it was also apparent that some trainers wished to increase their influence over the process, with issues such as initiating WPBA, choosing cases and timing assessments being seen by some interviewees as more appropriately trainer-led. Those who used WPBA were likely to be very aware of issues and problems in its day to day realisation, and may have felt they had good reason to adapt central policy regarding the assessment process to local circumstances.

Some of these issues will be further discussed in the chapter which follows, which addresses interviewees' opinions of the utility of WPBA, the perceived weaknesses of the process and the improvements they would suggest.

8: Theme 4:

WPBA's utility, weaknesses and suggested improvements

This chapter is divided into three parts, the first (sections a-c) concerning interviewees' perceptions regarding the utility of WPBA, the second (d) users' opinions of any weaknesses of the process and the third (e-f) their suggestions of how it might be improved. Although items included in this chapter have some overlap with material presented within other themes, such as users' views regarding who should initiate assessments, there is a sharper focus upon the perceived weaknesses of WPBA, and how they may be addressed, due to the evaluative nature of this chapter.

a. The utility of WPBA in assessing complex work

This section addresses whether users feel that the assessments are able to address all aspects of the radiological work for which they were intended. This stems from Grant et al's (2007) evaluation of the use of WPBA in the Foundation Programme, which suggested that the assessments were only appropriate at lower stages of training, and less well suited to the evaluation of complex work.

Training in clinical radiology, in common with all specialty training, represents progression beyond the Foundation Programme, and trainees are likely to be involved in more complex work from an early stage of their radiology career. It is important that the WPBAs they will undertake are able to address this work, and parts of the interviews

were specifically used to seek participants' views regarding the utility of the assessments in the evaluation of complex work.

Assessment of complex work does not represent the only facet of WPBA's perceived utility one could include, and it is followed by a section (b) in which users' opinions regarding the utility of the process are compared with previous means of assessment in the workplace, such as an overall consultant comment at the end of a period of training.

Following a review of how interviewees might view the term complex work when applied to radiology, this section explores trainers' and trainees' views of WPBA's utility in assessing both the technical steps and associated behaviours displayed whilst undertaking a procedure. This is followed by an analysis of participants' views regarding the difficulty of assessing subtle, nuanced work, and the practical problems which might be encountered when trying to record the results.

The interview questions regarding WPBA's utility in assessing complex work were subdivided between enquiries regarding assessment of knowledge and technical skills, and of behaviours, such as leading a team or dealing with an unexpected event. This division reflected Poikela's (2004) view, raised in the introductory chapter, regarding traditional forms of assessments' weakness in assessing behaviours, and it was important to assess if users felt whether WPBA could address this.

Although it is likely that the interpretational and procedural work undertaken during radiology training is generally more complex than that which trainees perform prior to entering the specialty, there is a very wide range of complexity of items upon which trainees may be assessed, ranging from interpreting a radiograph to performing a multi

stage procedure. The latter might include explaining the procedure to the patient, obtaining informed consent, then undertaking a long and complex examination requiring much planning, thought and manual dexterity in order to execute it.

When discussing complex work, it generally felt as if those I was interviewing thought the questions concerned high-end (such as complicated, multistage practical procedures) rather than mid-range or low-end radiological examinations (such as less complicated procedures or interpretational work).

Although placing activity into these categories is inevitably subjective, it is important that enquiry into complex work is not regarded as a simple proxy for all activity in radiology. It felt important that I took account of this when analysing the data, realising that respondents were likely to be discussing a subsection of radiological activity and that their answers might not apply to all work trainees might undertake.

Despite the potential uncertainty of what might be regarded as complex, the interview query still felt like a useful means of enquiry into whether WPBA represented a valid means of evaluating radiological activity and determining which (if any) aspects of this work were well suited to this means of assessment. I thus sought an indication of whether users felt WPBA could satisfactorily reflect trainees' actual performance.

Attempting to enquire about WPBA's utility in the assessment of complex work did reveal some commonly held views amongst respondents, but there were also some areas of little consensus. There were members of both groups who felt that WPBA could be applied to complex work, and half of the trainees thought it would be possible;

'I think the discussion that you have after or whilst performing a workplace based assessment enables you to break down the complex procedures into steps and analyse them and see how you've done from the beginning to the end.' (Trainee 7)

and

'I think there's still plenty of scope within the workplace-based assessment to cover all the important parts of it. So, yes, I think they're equally as valid for complex cases as they are for simple ones'. (Trainee 4)

There was support for this view from a similar proportion of trainers, two examples, being as follows;

'I've been involved in writing a DOPS for oesophageal stents for the BSIR and, actually, when I got involved with that, I did think at the start, "Well, there's absolutely no chance you could really do this very well," and, actually, the document has been produced, has been seen by a couple of other people and they said it was brilliant.' (Trainer 1)

and

'You'd have to break it down into compartments, and simple one that I use is, have you got the right patient, are we doing the right thing, is it safe to do if you're looking at a practical procedure? Obviously that needs to be expanding into not only technical ability of doing the procedure but also the post-procedure care. But anything complex can be put into workplace assessment.' (Trainer 7)

These quotations illustrate quite a mechanical view of assessing a complex procedure, where it is broken down into individual steps, each of which may be observed, fed back upon and signed off. However, aside from the second trainer quotation, there was less mention of taking a holistic view of such work (as discussed in the literature review) and ensuring that the technical steps were formed into a coherent overall performance.

It might be difficult for a trainee to perform a procedure completely incoherently, but it is still important that an assessor takes an overview of such work to ensure that the trainee works in a correct sequential fashion. Good performance of a procedure might be seen as greater than just fulfilling all the small steps it comprises, and it is important that a holistic view forms part of the assessment.

Trainers and trainees who were enthusiastic regarding WPBA's ability to assess complex work generally commented in similar terms to each other, but there was less uniformity regarding its perceived utility in assessing behaviours such as leadership, team working or dealing with an unexpected event.

In contrast to the positive opinions quoted above, other members of both groups were more sceptical of WPBA's ability to assess complex work and such negative views were articulated by trainees, who said;

'It's more of a general overview of how you approach the patient and how your needle skills are and things like that. I think it's more of a basic level I would have thought, not for a complex intervention.' (Trainee 6)

and

'When you're doing something complex there are often multiple aspects to it, which you have to address in a stepwise fashion. And you might get certain aspects of it done very well, but the overall picture might not be exactly what you're looking for. I'm not sure the workplace based assessments allow you to point out that you've done particularly well in one area, but there's a certain facet within a case or investigation or examination which you need to improve on.' (Trainee 5)

The comments are again mechanical in nature, reflecting the perceived difficulty in observing all the steps in a complex procedure, and concern that detailed feedback might

be subsumed into a general evaluation of a trainee's work. In addition, the first quote also suggests agreement with Grant et al's (2007) view that WPBA was only suited to assessing more basic work. A similar view is expressed in the first quotation below, and two trainers, who were sceptical about WPBA's ability to address complex work, expressed their views in the following terms;

'I think, if you're talking about complex work-based activities, I think there is still some scope for this process, for example, at doing practical procedures, some of which may be quite complex. There is still some scope for interpretation of somewhat more complex cases. But I think if you're thinking about more complex work-based processes, such as, for example, how does the individual interact with his colleagues or her colleagues, other behavioural aspects, I think that's much more difficult to capture in an exercise like this.' (Trainer 2)

and

'One of the things which is lacking in assessment is does the trainee know when to stop the case, when to accept the result? It's not about knowing the limitations, that's different. It's where the trainee knows he can't do it any more himself. But in complex cases you have to accept certain amounts of less than satisfactory results and you have to say "Fine, I'm going to accept this." So there are those kinds of complex issues; which I don't think workplace-based assessments on a tick box system gives you that.' (Trainer 6)

These quotations question the ability of WPBA to assess complex work, particularly in the domain of assessing non technical attributes, and both interviewees volunteered information regarding perceived deficiencies in the assessment of behaviours and professional skills, such as the fine judgement needed in order to stop a case and accept a suboptimal outcome. Some trainees expressed similar views, and even those who were enthusiastic about WPBA's ability to capture the technical aspects of a procedure expressed doubt about its ability to assess behaviours, as in this quotation;

'Assessing your general overall team working like that is more difficult using the workplace based assessments. I think that's somewhere where prose or asking someone's opinion independently is probably a better assessment.' (Trainee 7)

There were individuals in both the trainer and trainee groups who felt that a form could be written for anything (and by inference, anything could be assessed), but the trainee who expressed the opinion quoted above also suggested that trying to assess team working and capture all the steps, nuances and behaviours present in a complex case would be practically very difficult, saying;

'You end up with a not a feasible means of assessment. On the forms that they initially brought out there were so many steps to rank people on that there's no way any assessor would want to sit down with a form that's four pages long.' (Trainee 7)

This quotation raises other important issues regarding aspects of WPBA's utility, these being practical constraints associated with the assessments such as feasibility and the amount of time needed to undertake them. As stated by the interviewee, it might be possible to produce a very detailed form to capture many aspects of a trainee's performance, but if it is not feasible to use it (or only feasible to use it superficially or incompletely) due to insufficient trainer time to complete it, then its utility is likely to be much reduced.

b. Comparison with end of attachment assessment

This section investigates the perceived utility of WPBA in comparison to the previous method of assessment in the workplace, which was a trainer's (or trainers') comments

upon the trainee's performance at the end of a period of training. This practice continues in tandem with WPBA, and to some extent, should be informed by the assessments.

Trainers passing overall comment at the end of a training period regarding trainees' general progress has been criticised in the past as having the potential to put trainees' progress at risk, because an individual trainer might fail them for incorrect reasons, such as a clash of personalities. The system was also criticised for lacking specific evidence upon which important decisions regarding a trainee's progress might be made, and WPBA was designed to address these perceived failings, by providing a record of assessed episodes witnessed by a variety of trainers.

Despite its potential problems, it was clear that some trainers and trainees held the final consultant comment at the end of an attachment in high regard, although it might be one person's, or a small group of people's, opinion. This was reflected by a significant minority from both groups placing greater emphasis upon the overall consultant comment at the end of an attachment rather than WPBA (even as a collection) as the optimum means of assessment, as demonstrated by the following quotations from two trainees;

'I feel that is the supervisor consultants; he is the one who decides if I pass or fail regardless of my... I can have six/seven work-based assessments that say I'm excellent. If he thinks I'm rubbish, I'm probably rubbish.' (Trainee 2)

and

'The most important thing to me is that, at the end of the three-month attachment, is that the consultant says to me and signs off my form to say that I've reached a level that they would expect, or if not above the level that they would expect, for somebody at my stage. And as long as I get that, to be honest, what I've got for the workplace-based assessments, I'm much less concerned about.' (Trainee 4)

A similar view was expressed by some trainers, who felt that the role of assessing trainees at the end of an attachment should not be based upon WPBA, two saying;

'I think the main thrust of it is teaching. The assessment, I think, is a secondary thing, as far as I see it, because I think we do other things which are more of an assessment, which are, in radiology, obviously exams, but also in radiology, the formal end of attachment appraisals and so on that we do with the trainees. I think those take more into account, you know, what we've felt in terms of the trainees' competence.' (Trainer 2)

and

'... The consultants on the team, to provide any insight that they can into the performance of the trainee, and that will be based on your observations of their day to day behaviour and performance in the workplace on which you will make judgements the whole time. It's not a formal assessment but I think overall no one's ever shown that any of these tools is actually any better than the opinion of multiple experienced trainers.' (Trainer 5)

Although interviewees were not directly asked regarding their opinion of other methods of assessment, they were asked how well they felt WPBA fitted in with other assessments and several respondents compared WPBA with what had gone before. Their answers suggested that (no interviewees described WPBA as better than or of equivalent utility to the end of attachment comment) utilisation of the results of WPBAs as indicators of a trainee's progress through an attachment appeared not to have been widely adopted. In addition, there was little evidence that the assessments' results were used to inform the end of attachment comment to any great extent, one trainer saying;

Q: But in the scheme of the attachment, in the great scheme of things, how much effect are they having?

A: I think small. (Trainer 9)

There is no evidence that WPBA was designed to supersede the system of consultant comment, and to some extent they address different aspects of training (episodes of formative assessment versus progress through an attachment) but it is apparent that a significant number of interviewees still viewed the trainer's final comment as the most important indicator of their day to day progress, and that the results of WPBA did not form a significant contribution to that judgement. As with the results of WPBA, the end of attachment comment also forms part of the training record, is recorded in the trainee's portfolio, and contributes to the judgement regarding progression made at trainees' annual review.

c. Summary of the utility of WPBA

The preceding sections have looked at users' views regarding the utility of WPBA, from the perspectives of whether it may be used to assess complex work and how it compares to the system of consultant comment at the end of a period of training.

There was little consensus either within each group or between trainers and trainees regarding the utility of WPBA for assessing complex work, with reasonably balanced numbers believing it was or wasn't possible. Amongst those who thought it was possible, there was more enthusiasm for WPBA's ability to assess technical steps in a procedure than for its suitability to address attributes such as behaviours or team working. Concerns were also expressed regarding practical issues such as long evaluation forms, which might not be completed due to pressure of time.

In terms of comparing WPBA's perceived utility with that of the overall consultant comment, it is perhaps unfair to contrast assessments of individual episodes which might inform an overall comment with the comment itself. However, a significant minority of trainers and trainees clearly valued the latter more highly, because of its wide ranging view regarding overall progress (rather than individual episodes of WPBA) whilst acknowledging that such comment was based upon opinion. Although not all of the interviewees discussed the consultant comment, no negative opinions were expressed by those who did.

The role of WPBA in contributing to the consultant comment seemed very slight, if the assessment results were used at all. This runs counter to one of the proclaimed advantages of WPBA, specifically the numerous assessors and assessments which would inform it, thereby negating the power of one trainer (or a small group) to pass a judgement (which could have little supporting evidence) which might block a trainee's progression.

d. Weaknesses of WPBA

This section explores users' opinions regarding some of the weaknesses of WPBA, in addition to some interviewees' reservations regarding the utility of the assessments described above. The areas of perceived weakness raised by some users were concerns that WPBA may represent a tick-box exercise, the impact of time pressure upon the assessments, its failure to identify poor performers, the use of generic rather than specific evaluation forms and trainers who were poorly prepared at the time the scheme commenced.

Some of the interviewees' perceived weaknesses of WPBA have inevitably been included in earlier chapters, but this section serves to unite them with others which haven't yet been raised, and the section after this explores the improvements respondents suggested, followed by a common summary.

The commonest weakness of the scheme, in the opinion of both trainers and trainees, was that it represented a tick box exercise. This phrase might be interpreted in two ways, neither complimentary. Firstly to describe a process with little perceived usefulness merely completed to ensure trainees passed their annual review, or alternatively, an attempt to adapt WPBA to a binary system of marking, allowing assessors to tick boxes when recording their views.

Viewing WPBA as a tick box exercise was also a frequent complaint amongst the psychiatric trainees surveyed by Menon et al in 2009 and the anaesthetists studied by Bindal et in 2013, reflecting a view that WPBA had few real benefits for training and was a peremptory exercise undertaken to fulfil the conditions of trainees' annual review.

A trainee interviewee expressed this in the following terms;

'I suspect that another way of abusing it would be a more low-level type abuse where maybe, if you had a consultant who was less interested in it, and was just willing to tick boxes to get it over and done with.' (Trainee 5)

Although only a small minority of trainees expressed a view of the WPBA scheme in these terms, a third of trainers interviewed did, with an example as follows;

'I think the bad thing, kind of, really emanates from the trainees who don't utilise it properly. So I think if the trainees are just seeing it as a, "Oh, can I please do five Mini Interpretation Exercises with you during this attachment?'

Let's just do them now. These are the cases that I've seen. Let's do it here. Let's do it now. Can you just tick this box?" That's not right. And I think that becomes meaningless, and I'm afraid to say that maybe the minority of trainees are seeing it in that way. And I would suspect some of the trainers are seeing it in that way, too.' (Trainer 2)

In these examples, and those cited in earlier sections, both trainers and trainees are implicated in treating the assessments in a peremptory fashion. This may be because some trainees see achieving numeric targets as the greatest priority, or even lack of commitment to the assessment scheme on the part of members of both groups. Another possible reason, cited by a minority of both groups, was pressure on both trainees' and trainers' time, this trainee saying;

'I think it's the schedule of the trainer and the trainees, and to get one time where you've got a good 15 to 20 minutes, depending upon what you are doing, to get that free time, it's very tough. I think that's one of the biggest issues with workplace assessments, as far as the trainee's concerned and, I think, as far as the trainer is concerned.' (Trainee 8)

This view was supported (overtly) by one of the trainers during the following exchange;

Q: 'Thinking about having done one and your general pattern of work, how feasible is it to do in your work day?'

A: 'I think it's very difficult. And I think if radiology is going to embrace this system then we have to find some time somehow, whether that is within our timetables or dedicated sessions.' (Trainer 12)

These quotations illustrate that issues of trainers' available time may lead to users having to undertake the assessments quickly, representing a potential problem (in some cases) with regard to the feasibility of undertaking WPBA.

The main additional weakness perceived by trainers was a failure of the scheme to identify poor performers, cited by a third of the trainers interviewed. This is particularly interesting, as it represents one of the specific functions some trainers thought WPBA should fulfil. Generally trainers did not state where they thought such identification should lead; although a trainer quoted upon this subject (who saw WPBA as having both formative and summative purposes) was unambiguous in stating his view;

'I think there's also a potential for individuals to pass Workplace Based Assessments who are hopeless, and there is a particular individual who's currently in the training scheme who is, in my opinion, a hopeless radiologist and, actually, I think, is a pretty hopeless doctor, with no insight and yet, despite the current system, we're unable to provide, or obtain, a level of proof that is sufficient for him to be removed from training.' (Trainer 1)

This quotation suggests that WPBA may have difficulty in addressing subtle (but very important) areas or wider issues such as trainees who are perceived to have a lack of insight into their problems, and this was supported by this trainer, who said;

'People are reluctant to put down criticism, and people may just still scrape by as satisfactory, and workplace based assessment is just one aspect of that person. A lot of the time the failings of trainees are to do with their insight, or... well that's one of the main worries, insight and maybe motivation and I think those wider things may not be assessed by that assessment.' (Trainer 4)

There are also issues regarding some trainees' selection of cases, as outlined in chapter six, and how choosing unchallenging subjects for assessment might allow a poorly performing trainee to conceal the situation. Although central guidelines regarding WPBA suggest that trainees choose assessments of different subjects within their

practice, there is no guidance regarding how complex such cases should be or any suggested proportion of complex and simpler cases.

The issue of choosing appropriately might be solved by either joint or trainer-led selection of cases for assessment and was advocated by interviewees from both groups. It is discussed further in the suggested improvements section.

Another weakness of the scheme identified by both trainers and trainees was the use of generic forms to record the results of a large variety of assessments. This led respondents to comment that many categories were not applicable, a trainee saying;

'Often there is a not applicable box that can be ticked, but I can't think of a specific example. But I do remember a couple of things that I've done where a large chunk of the form seemed to be not really applicable.' (Trainee 5)

Issues such as this could be seen as less important than those cited previously, as there are areas on the forms for trainers to enter free text, although the potential utilisation of more specific forms was mentioned by two trainees, and is included in the next section.

A more serious weakness of the scheme highlighted by two trainees and two trainers was that trainers appeared unprepared for the assessments when the scheme commenced in 2010, and as a result had to learn how to utilise the system with trainees who may have appeared more computer literate. This situation was outlined by a trainee and a trainer as follows;

'I did a DOPS. It was the first time that this particular consultant had used the ePortfolio, so there was a little bit of stress there from her point of view, so we... I sent her the tickets to do the DOPS and we sat down and helped her log in, and explained it to her, explained the system.' (Trainee 1)

and

'I think the way I became familiar with them was looking at some of the web-based videos and so on, which are available on the RCR website, showing how to use these tools. And also, with the trainees, initially, we were almost learning together how to use them.' (Trainer 2)

These quotations suggest that some trainers were not well prepared for WPBA when the scheme was introduced and to address this, some sought information on the internet, some taught themselves, and some learnt the system by working alongside the trainees they were assessing.

Although this suggests that some trainers had not been trained in the use of WPBA when it commenced, by the time the interviews were conducted (starting 6 months after the scheme's introduction), two thirds of the trainers interviewed had received some form of training (either supervisors course or local lecture) in undertaking WPBA.

However, the number of trainers in the sample who had been trained to undertake the assessments may not have reflected the proportion of those trained in general, particularly if the respondents (as might be the case) represented those who were more enthusiastic about training. Worryingly, a third of the trainers interviewed said they had not been taught how to undertake the assessments, as evidenced by the following exchange;

Q: 'Were you trained in how to use the tools?'

A: 'No.'

Q: 'How did you feel about that?'

A: *'I think it's typical of the current culture, we just get just expected to do things.'* (Trainer 11)

Such trainers will have undertaken the assessments without any formal training, and although there were some trainers who had sought information regarding WPBA elsewhere (e.g. the college website, as per the second of the paired quotations cited above) others might have conducted assessments according to their own view of best practice.

It should also be noted that there was not specific training in WPBA for trainees, although they may have had experience in undertaking similar assessments prior to entering radiology, and they had the same access as trainers to resources such as the college website.

Though training trainers in the use of the assessment tools may not have served to ensure they were used according to central guidelines (trainers may still have adapted or sought to improve WPBA for reasons they felt legitimate) it might have helped introduce some uniformity into their enactment. Training users is one of the topics covered in the next section which concerns interviewees' suggestions regarding how the assessments might be improved.

e. Suggested Improvements

In this section, interviewees' suggestions with regard to addressing the weaknesses identified in the previous section are explored, as are other improvements suggested by

respondents. These include seeking to reduce the tick box nature of some episodes, greater training in WPBA, joint or trainer leadership of the process, random case selection, the timing of WPBA during an attachment and refining the assessments themselves.

In the previous section, the negative aspect of WPBA most commonly cited by interviewees was that the assessments represented a peremptory, tick box exercise for some users. However, possible solutions to this problem were offered far less frequently, although one trainer suggested;

'It's hearts and minds. So there's an educational process, there's the redesign of the forms to take away the tick box element so you have to give feedback. There is a training element; I don't think that everyone who does them is trained.' (Trainer 5)

The quotation has resonance with the closing paragraphs of the previous section, in suggesting that greater training of users will lead to their using the assessments in a formative fashion, and also acknowledges the influence of the score sheets in the process (where the boxes are actually ticked).

There is the suggestion (as above) that redesigning the forms would help, the inference being that removing tick boxes might leave more room for results to be expressed as a narrative, aiding feedback delivery. Such redesign might also help to remove any element of grading or mechanistic criterion based assessment from the process, and the potential influence of both these factors upon assessment was discussed in the literature review. The tick boxes relate to particular aspects of performance, and their removal

might help discourage grade seeking behaviours from trainees and encourage assessors to adopt a holistic view of trainees' performance.

The overarching messages from the quotation are a perceived need to emphasise the formative nature of WPBA by training users and removing the drivers which encourage peremptory assessments. Removing numeric targets might address the latter, but without them some trainees might not engage in the process.

It should also be noted that despite the suggestion that greater training might encourage more formative use of WPBA, material quoted in the literature review (Gingerich et al, 2011 and Govaerts et al, 2013) suggested that such training might not be as effective as expected due to assessors' tendency to form categorical judgements and utilise personal epistemologies when undertaking assessments.

Rather than attempting to ensure all assessors adhere to central guidelines regarding the assessments' use, which has generally been the strategy utilised at RCR training events I have attended, alternative approaches such as allowing groups of assessors to define performance levels themselves within frames of reference (Govaerts et al, 2013) might lead to greater assessment consistency and buy-in on the part of trainers.

The practical issue of time pressure was widely cited as a problem, particularly by trainers, and several suggested the need for protected time to undertake the assessments, this trainer saying;

'There has to be suitable time allocated to do this but we are a training centre and I think that is taken on board anyway. So it does take a little bit more time. To be honest, I think if there was more time there probably would be more DOPS assessments.' (Trainer 10)

Both this trainer and others clearly found that there were conflicting pressures between the educational (including WPBA) and other aspects of their work, particularly service delivery.

Several interviewees suggested that the WPBA scheme could be improved by joint or trainer leadership, and that if this occurred, trainees would be able to pre plan an assessment programme with their supervisor at the start of an attachment. This was advocated by both groups, a trainee saying;

'I think that it would probably be an easier process if the trainers were slightly more interested from their point of view to say, while you're here in your three months we will make sure we do this many assessments from the word go, so that you don't mind asking them. If that's at an outset part of your set up is, right, we're going to do 10 procedural assessments or IPXs while you're here in three months, you must ask me once a week to do this, that would be useful.'
(Trainee 7)

Some trainers expressed similar views, and were keen that they should have an input into choosing appropriate cases for assessment, these trainers saying;

'I think the consultant should have some input in many ways. You don't want them to pick something so easy that you can't really assess them on it. So I think the level of difficulty... I think we have the experience to work out where something's appropriate.' (Trainer 10)

and

'I like to choose the case because I think, given my expertise and my experience, I'm probably in a slightly better position than the trainee to select cases which may have a teaching point behind them.' (Trainer 2)

As the above quotations suggest, one of the perceived advantages for this shift in emphasis is trainers having the experience to choose, and knowledge of, the most appropriate teaching cases with the additional advantages of avoiding easy cases or those chosen in retrospect.

These quotations show that representatives of both groups thought that joint leadership might improve both the organisation of, and learning from, WPBA, whilst another trainer cited an additional reason for increasing her input, saying;

'This is one of the reasons that I wouldn't want the trainee to just choose their own cases. I mean, with that particular trainee we had several workplace based assessments done by not just me but some of my other colleagues, and the... you know, there didn't appear to be any concerns on those particular assessments. But speaking to them all separately they all had concerns about this individual's overall performance.' (Trainer 3)

This demonstrates another reason for trainers wanting greater input into case selection for WPBA (and also illustrates trainers' perceived need to coordinate a series of results), in order that those that they suspect of being poor performers cannot conceal themselves by their selection of unchallenging cases for assessment by different assessors.

Revealing trainees who are performing poorly also concerns how discriminatory the assessments are in terms of both identifying underperformers and those who are doing well. Greater trainer input into selecting cases and coordinating the results might help in this regard, but an additional factor is discussed by Crossley et al (2011) in a paper looking at the nature of the evaluation scales used in WPBA.

They found that such scales are more likely to discriminate between trainees performing well and poorly when the ratings are aligned with constructs of increasing trainee independence and the development of clinical sophistication, rather than rating performance above or below expectations, as occurs now. This might be seen as a case of preferring criteria referencing to norm referencing, and suggests that improving the discrimination of WPBA depends on refining the design of the assessments themselves, as well as amending how they are used.

An alternative method of case selection advocated by a minority of trainees and trainers was that cases be selected at random, which might avoid selection of cases known to be unchallenging or those which had (in retrospect) gone well, and also impart a sense of fairness to the process. This was advocated in the following quotations;

'I think it's nice to approach the other consultants and say why don't you choose random cases where the trainees have no idea that you're going to assess them. And the trainers come and tell them, "We're going to do this now. Have a look at this case. This is a case we're going to do as a work-based assessment" – I think that would be very effective.' (Trainee 2)

and

'Because I feel that that would lead to a greater sense of fairness on the part of the trainee ... it's also quite difficult to know with a case until you get into it what it's like, whether it's suitable. But if it was an understood system that is random, and the trainee and the trainer have no part in choosing that, then I think that's appropriate in terms of cognitive radiology.' (Trainer 11)

It should be noted that none of the trainee interviewees alleged that cases selected by trainers were seen as unfair, but trainer input and/or random case allocation were seen as ways of trying to avoid inappropriate or retrospective case selection. However, if cases

were chosen at random, there would be a significant probability that normal or unchallenging examinations might be selected, whilst those with particular teaching value might not be utilised.

Another benefit of trainers and trainees jointly leading WPBA, as illustrated by the earlier quotation from trainee 7, and supported by trainers quoted below, would be spreading the assessments through an attachment, using them to aid trainees' development week by week. Sometimes assessments are left until the end of attachments, when the opportunities to improve in a particular sub specialty might be very time limited. This aspect was highlighted by these trainers;

'I don't think there is a culture yet of let's do this every week for the whole 12-week attachment which is, as I understand it, the way it's meant to be. It's not meant to be near the end of the attachment, it's meant to be all the way through and used as a tool of how can we improve on this each week.' (Trainer 4)

and

'I think ideally you would perform them at multiple stages during the attachment, and I have been relatively so far not very proactive in organising them, so letting the trainee request the assessments. They tend to request them at the end of the attachment.' (Trainer 9)

Other suggestions for improving WPBA concerned the form for recording the results, which was found too generic by some interviewees, who advocated the use of greater numbers of forms specific to particular activities, as suggested by this trainee;

'I think that there are many subspecialties and specialties such as (interventional) radiology and it's difficult to maybe get your subspecialty interest to fit into some of the more generic forms.' (Trainee 3)

Aside from the more technical issue regarding the forms' content, it is also important to remember that the form may influence the conduct of the assessment, (as noted in the first quotation in this section) by containing criteria and rating scales which act as cues for assessors. Thus although rewriting them might improve the specificity of WPBA in certain situations, other factors (such as including or removing rating scales) might have a more general impact upon the assessments' ethos and conduct.

f. Summary of weaknesses and suggested improvements

Aside from the reservations raised in section one regarding the ability of WPBA to evaluate complex radiological work (and this may apply to other specialties too) the perceived weaknesses of the assessments described were generic in nature, and many were common to quoted experience from other specialties.

An example of this is the view held by some respondents that some trainers were poorly prepared for WPBA's introduction, which may have been partly due to lack of time to become familiar with the scheme. This perceived lack of preparedness may have played a part in some trainers' poor engagement and undertaking peremptory assessments; although some trainers felt that the latter was a trainee issue too. The suggested solution of protected time for trainers to train might help increase trainer availability, but service pressures mean that this may not occur.

Amending the initiation and leadership of WPBA was desired by representatives of both groups. Some interviewees felt it might address many issues they raised, including

inappropriate case selection by trainees, spacing assessments through a training period, and identifying poor performers by trainers choosing challenging cases.

Another suggested change which may be feasible is to produce more specific evaluation forms for WPBA in radiology, though this would need to be instituted with care, in order that comprehensive but unwieldy forms (which may not be filled in by users) did not result.

Redesigning the forms was also suggested as a means to help reduce the peremptory, tick-box nature of some assessments, although the forms may just be a record (or an exacerbating factor) of such practice rather than its root cause. Removing an annual numeric target might also reduce the number of peremptory assessments, but a likely undesirable effect of this (as noted by a trainee in an earlier chapter) would be a reduction in the numbers of trainees engaging in WPBA, and a likely fall in the number of cases each had assessed.

9. Discussion

This chapter serves to return to some of the subjects introduced during the literature review and developed in the data analysis. Some topics such as the purpose of WPBA reside in both the literature review and data analysis, and are further addressed here. Others, such as how users adapt the assessments for day to day utilisation, arise from the data analysis. Adaptation might be seen as a development of theme three (how users influence the realisation of WPBA), but in addition to this, the discussion also aims to place such adaptation in a wider context by comparing it with examples from other professions and investigating why it occurs.

In addition, this chapter aims to view material from the data analysis from a different perspective, by collating information from across the themes and conceptualising it differently. This informs sections which discuss power and control in WPBA (both on organisational and individual levels) and why respondents may indulge in activities such as gaming and subversion of the assessment process.

As was the case with the thematic analysis of the data, there will inevitably be overlap between these sections and difficulty demarcating certain areas, an example from the discussion being users' adaptation and subversion of the assessments. Whether an action falls under one or other category is subject to the author's judgement and factors pertinent to one area will often influence another.

a. The purpose of WPBA

In terms of evaluating the perceived purpose of WPBA within the medical education literature, it is difficult to find much debate regarding whether the assessments are for formative or summative use, with many authors appearing to regard their developmental use as established. However, this is not unanimous, as illustrated by Al-Kadri et al (2013) who found that supervisors of medical students in their institution tended to view such assessments as summative in nature, without any additional role in teaching and learning.

The RCR's policy (as noted at the start of chapter 5) regards WPBA as a formative process, and when the assessments' purpose was discussed with interviewees, their opinion regarding its formative or summative purpose formed the basis of the discussion. The division between the two was discussed by all interviewees, and all other perceived purposes, such as identifying poor performers, were only raised by a minority.

From the data analysis, it is clear that both trainees and trainers see WPBA as primarily formative, with the delivery of feedback as an important part of the process. However, there is a divergence in the two groups' views with respect to its summative usage and more trainers than trainees appeared enthusiastic to utilise the assessments in a summative fashion, generally using the collected results of numerous episodes.

Some of the differences between the two groups, such as interest and disinterest in identifying underperformers, might be explained by the differing perspectives of trainers

and trainees, as a trainee need only be concerned with their own progress. Trainers may be responsible for many trainees and see it as their role to uphold standards of clinical work, whilst identifying those who need assistance to improve.

It is clear that trainers and trainees may have varying expectations of WPBA (although with much overlap between the two groups' views) and this might be expected due to their differing perspectives as noted above. These varying expectations may lead to issues such as whether a trainee considers an assessment fair or not, and are further explored in the paragraphs which follow.

The reasons for some assessors adapting the purpose of the assessments are likely to be multifactorial, but it should be noted that many trainers will have become used to making judgements about trainees' progress on an informal basis in the past, by means such as commenting upon work during an attachment. They may be unable to completely divorce this from their participation in WPBA, and thus may use the new system in a more summative fashion than its designers (and the college curriculum) intended.

Part of a consultant assessor's clinical role is likely to comprise forming judgements about patients' imaging and producing definitive reports about it, and some may carry this over into their work in training. This judgemental role with respect to training might be expected to gradually change as those trained and assessed by WPBA become trainers themselves.

The difference between trainees' and trainers' expectations of WPBA is also influenced by the authority relationship between the two groups. Whatever purpose is perceived by

the trainee, once the assessment has been initiated and the assessor chosen, it is likely to be the latter's view which prevails, as it is they who have charge of the assessment, record the results and deliver feedback. Issues of power and control within WPBA are further pursued in a subsequent section.

The potential importance of divergence of a trainer's and trainee's view of an assessment's purpose is that it may disadvantage a trainee who views it as a low-stakes formative episode, prepares accordingly, and then finds that the stakes may be higher than expected. Ambiguity of purpose is also likely to disadvantage trainees in terms of perceived unfairness and the potential for the results to be extrapolated beyond their expected usage.

In addition, the results form part of a trainee's record, and for these reasons it is essential that there is clarity regarding both the purpose and conduct of WPBA, and that trainees know when they are being assessed, in order that they do not feel they are being treated unfairly. Although it may be unrealistic to expect an assessor to remove all thoughts of future summative judgement from their mind whilst conducting WPBA, the college curriculum states that single assessments should be judged (and fed back) as formative episodes, in line with the expectations of the trainees who undertake them.

The importance of treating single episodes in this fashion is supported by material from the literature review where authors such as Norcini (2007) emphasise the need for multiple assessments prior to any generalised estimate of performance. In addition, Govaerts and Van der Vleuten's (2013) suggestion that both competence and

performance are dynamic, context-related entities, subject to much variation support the contention that single assessments should not be used to make a summative judgement.

Outside the discussion of formative or summative usage, purpose could also be seen by some respondents in more mechanistic terms. Thus it was clear that some individuals from both groups saw it as a means of recording, documenting and auditing training, rather than primarily a developmental or final assessment process.

In some interviewees' opinions, the documentation of training and formalising its recording was of prime importance. This demonstrated a practical view of the purpose of WPBA in terms of its day to day utility, and their responses are perhaps best seen as pragmatic opinions regarding the day to day usage of WPBA, rather than a view of the assessments as developmental or final.

It should also be noted that the perceived formalisation and the imposition of a framework upon training suggested as a good feature of WPBA by some interviewees may be only partly due to the assessments' introduction. It is important to realise that training previously existed within a formal framework, most of which remained in place when WPBA was introduced. Although WPBA may have altered assessment of trainees in the workplace and given users an impression of different organisation, perhaps due to the documentation of the assessments, WPBA is still part of this overall framework.

A new electronic trainee portfolio and a scheme of educational supervision were introduced at the same time as WPBA, with the portfolio acting as a very comprehensive repository within which all aspects of a trainee's progress are recorded. These include the results of all assessments, their trainers' opinion of their work during

attachments, the outcome of each annual review of progress, research, audit and teaching activity. The portfolio's use is compulsory and universal and it is likely to have played a large role in making training feel structured, as well as that attributed to the assessments.

In addition to the perceived purposes already discussed, other interpretations of the process came in to play, such as needing to meet numeric targets of WPBAs to ensure progress through training. Unsurprisingly, the majority of trainees interviewed did not admit to such a view of the assessments' purpose, although quotations from chapter five show that more than one saw it in those terms and just did the minimum number required to progress.

This might be seen as a mechanistic, responsive view in the face of a mandatory numeric target. Achieving such a target (and no more) with the aim of ensuring progress in training might be seen a manifestation of gaming or playing the system, as one would hope that if WPBA were truly being used formatively, a trainee would aim to undertake as many assessments as possible in order to improve as much as they could.

To put such a target in context, the minimum annual numbers of WPBAs stipulated by the RCR's curriculum is six of each assessment (mini-IPX and Rad-Dops) per year, an average of one per month. In an analysis of national usage for a year commencing in August 2011 (Booth, 2012) it was shown that the mean number of assessments undertaken per trainee exceeded ten for each type, suggesting that many trainees exceed the minimum target and may thus be embracing WPBA's formative ethos.

My personal experience of trainees who are progressing well is that they usually exceed the annual target of WPBAs, whilst those whose progress is more marginal may struggle to achieve the numbers required. In some cases, this may indicate poor engagement with the training process and low numbers of assessments may be an important factor in identifying trainees who are struggling, aside from the results of the assessments themselves.

However, the study data also demonstrated that there were trainees who engaged 'just enough' with the process to get by, even though all other measures of their progress were above average, demonstrating that low (but above the annual target) WPBA numbers did not always reflect poor performance overall. Both interviewees falling into this category in the study were senior trainees who were already in training when WPBA commenced in radiology, and both intimated that they were being well taught and the assessments' introduction did not improve this. The proportion of such trainees might be expected to fall in future, as trainees will have engaged with WPBA from the start of their radiology training, and will not have been used to a different method of learning and assessment which they might want to continue.

Playford et al (2013) found an alternative pattern of trainee engagement in a study of undergraduate WPBA which demonstrated above target assessment numbers in the vast majority of a large number of students, regardless of whether they were weak or strong academically. The authors felt that this resulted from a supportive culture associated with a positive experience of appraisal. However, they also admitted that a summative element to the assessments might drive students to do more to increase their chance of

engaging in a high scoring episode whose result would be recorded (the best results from each discipline were added to the student's academic record).

Although WPBA in radiology has a formative ethos, material from other professions cited in the literature review (e.g. Stanley et al's (2002) paper regarding assessment of music performance and Gonczy's (1994) review of solicitors' assessments) demonstrate that WPBA may be used both formatively and summatively outside postgraduate medical training.

The way in which workplace (or context) based assessment is used outside medicine depends upon the context and the intentions of those setting the assessments. Even in the examples from outside medical education, the assessments' purpose appears to have been set (rather than debated) in order to address the situation in which they are used.

There is less data from other contexts regarding other perceived purposes of WPBA, such as achieving numeric targets, but the purpose seen by some interviewees of organising training is reflected in other contexts. In an analysis of the utilisation of performance-based assessment in mathematics teaching, Firestone et al (1998) argue that assessment policies; *'do more to organize existing learning opportunities for teachers than to increase them'*.

It should also be noted that the overall system of professional assessment in which WPBA is deployed may also affect its perceived purpose. In radiology (as in most medical specialties) there is a long established system of external examinations which are used for summative assessment, as outlined in appendix f. With summative assessment covered by another method enacted outside the workplace, there is a natural

opening for the formative use of WPBA. In other contexts, such as might pertain in the example of Australian solicitors (Gonczi, 1994) quoted in the literature review; there may not be another system of assessment, meaning that the competency-based method must be used summatively.

b. How WPBA is adapted for day-to-day use

The material which follows seeks to situate users' adaptation of WPBA within the wider contexts of both postgraduate medical training and examples from other professions. Included within the latter, is discussion of how inevitable (or not) local adaptation is when central guidelines are put into practice and why such amendment may occur.

It is clear from the data analysis that central guidelines for the conduct of WPBA are adapted by both trainers and trainees when the assessments are realised in the workplace. The range of adaptations of the process is covered in chapter 7, and conscious adaptation of the conduct of WPBA appears to be generally instigated by trainers, rather than trainees, which is not surprising, when the authority relations pertaining to the assessments are considered.

Adaptation tended to occur when users made an active decision to perform an action to amend the process or fill in a gap in the WPBA guidelines, rather than leaving an action unperformed. As inferred by the term 'gap', it must be acknowledged that guidelines cannot prescribe in all circumstances, and there may have been some instances where assessors had to make a judgement to enable an assessment to continue. In circumstances where there was not a gap and central guidance was not followed, the part

of it which was set aside was replaced by another action, such as trainers deciding to assess trainees without their prior knowledge or recording a different written account of an assessment than was given orally. There is further discussion regarding the limits of central prescription later in this section.

When discussing the implementation of central policy, O'Connor et al (2012), suggest that the most profound influence is at the; '*clinical coal face*' (i.e. the end users of WPBA) and that users may derail policies by either passive or active resistance. The former (ignoring policy) may not apply to realising assessment in the workplace, but amending policy can affect WPBA's enactment, and represents an active process, although perhaps not overt resistance.

The same authors make the further point that medical staff (who will represent the vast majority of assessors) tend to value autonomy, may be resistant to change, and may exclude themselves from policy implementation. This has resonance with interviewees alleging that some trainers do not engage with the process, whilst others adapt it in ways they see fit, with the potential to impact upon the fairness of the process as illustrated in the quotation below.

Despite this discussion of the need for fairness and uniformity in the process in this and the previous section, it is likely that some adaptation is necessary for WPBA to be realised satisfactorily in certain contexts, particularly if there are limitations on personnel or resources which may potentially affect the assessments. Outside a medical context, the need to adapt assessments to evaluate learners undertaking highly

contextualised work-based learning is addressed by Costley (2007), who argues that different assessment strategies are necessary to address differing approaches to learning.

Such opinions, and those quoted in the literature review, demonstrate how central policy may be amended on the way to (and by) end users, and that there may be a need to adapt WPBA when it is realised, in order to take account of local circumstances and meet learners' needs. However, such amendment may not always be inevitable, or made for altruistic reasons and there are potentially negative effects upon the fairness of an assessment, as illustrated by this quotation;

'But the consultant was sat there and I had to sit on the side of the consultant and he said, can you spot it, and because the controls were so different I just couldn't change the windows properly, and anyway I didn't get the diagnosis at the end. I struggled a bit with that, and he said, oh let's do an IPX on that, so I said that wasn't fair.' (Trainee 6)

Beneath the adaptation of assessment policy in radiology lies the question of why users seek to adapt WPBA.

It should be noted that in some cases it may be due to an altruistic desire to improve assessment, as suggested by Long (2001) and discussed in the literature review, who used her position as director of a local radiology training programme to make modifications to a previous RCR trainee assessment form in order to improve its utility.

However, there are other reasons that practitioners adapt assessment systems, including resistance to change and weak central guidance, as outlined in the literature review. If some assessors view the assessments as more summative in nature and wish to identify

poorly performing trainees (as stated by some interviewees) this may provide a motive for their adapting WPBA to suit their own purposes. As described above, medical staff tend to value autonomy, and may feel that they know how best to assess trainees, adapting central guidelines to suit their own practice and their local circumstances.

It is sometimes difficult to judge the significance of adaptations to the realisation of WPBA and there may be circumstances where different approaches might bring benefits to episodes of assessment and enhance their teaching value. It is important to realise that these may be instituted for positive reasons, with trainers aiming to support and encourage the trainee and improve the process. Examples of this might include choosing a case in retrospect due to its excellent teaching value, reporting an assessment in an overtly positive way to increase a trainee's confidence or only assessing part of a procedure were it known in advance that it represented a specific area where a trainee needed to learn and improve.

Some of these examples might be seen as assessors exercising their prerogative to bestow credit upon trainees (Sadler, 2009) as was discussed in the literature review. Other instances where this might occur and have a positive influence on the trainee and their assessment could include acknowledging instances of novel thinking, giving tangible praise for great effort or taking a holistic view of performance beyond the usual bounds of the assessment form. All of these might be viewed as assessors 'breaking the rules', yet could also be seen as having the potential to improve the assessment process.

Such an approach might be more easily instituted in a context such as higher education where Sambell and Hubbard (2004) suggest the role of formative assessment is to

support students' learning and help them realise their potential. Were this all that was at stake in postgraduate medical education and WPBA truly represented low-stakes formative assessments which weren't recorded, had no long term implications and changes were agreed in advance by both parties it might matter less what format was adopted during the episode.

However, it is likely that some of the originators of WPBA policy might argue that major variation in assessment format should not occur in postgraduate medical education, where trainees have to study and fulfil a national curriculum; the results of WPBAs are recorded and may be seen by third parties or affect future progression. In addition, it might be argued that the subject of WPBA (patient care) would be regarded as high-stakes and thus less amenable to variations in approach which might potentially compromise the rigour of the assessment.

c. Power and control in WPBA

Issues of power and control in WPBA have been mentioned briefly in previous sections of the discussion, and in this section the subject is further developed in respect of relationships between professional bodies and practitioners, and between assessors and trainees, when assessments are undertaken. The majority of this section concerns the latter, with particular reference to the factors affecting authority relations between trainers and trainees during the assessments, and how these might change.

The RCR issues central guidelines for the conduct of WPBA, but despite this, some trainer interviewees freely admitted to adapting the assessments, none mentioning any

anxiety regarding such amendment, and none volunteering any concern regarding whether or not such adaptation breached central policy. It must be admitted (as stated in the previous section) that the central guidelines cannot prescribe every action, or for every circumstance during WPBA, but there were instances where adaptation did not serve to just fill a gap in the guidance.

When I discussed the conduct of the assessments with assessors, some stated that they had received training in the process, which might be expected to help standardise practice, but even in those instances, there appeared to be a generally held view that assessors had a degree of autonomy, and none of those who adapted the assessments appeared to feel constrained by the influence of central guidelines. This aligns with the view expressed by O'Connor et al (2012) in the previous section regarding the value placed on autonomy by medical staff, their resistance to change and how some may exclude themselves from policy implementation.

Thus the degree of authority held centrally did not appear enough to inhibit some assessors from adapting the assessments, and it is difficult to see how such behaviour could be altered or stopped, unless trainees actually complained about the conduct of an assessment or an individual trainer's practice was highlighted in some other way. In the study, only one trainee mentioned an assessment episode which they felt was unfair, and I do not think they reported this outside of the interview. In addition, as trainees can choose their assessor, they could avoid (or even seek out) assessors who worked outside central guidelines and thus such practice might not come to light.

The comments in the above paragraph also allude to issues of power and control between individuals during episodes of WPBA. Interviewees (both trainers and trainees) discussed such issues far more than their relationship with central guidelines during the study. The reasons for this are likely to be twofold, firstly that individual assessment episodes constituted by far the greatest amount of interviewees' experience of WPBA, and secondly their interest at this level stimulated the interviewer's probes and questions.

Control of an episode of WPBA might be thought to be held by the assessor, as it is they who run the assessment, decide its outcome and record the results. However, the balance of power surrounding an assessment episode as a whole may not be as one sided, as WPBA is mandated as trainee-led, and it is their role (according to central guidelines) to initiate the process, choose the case for assessment and the assessor. Thus the balance may shift between trainee and assessor between the initiation of the assessment and its enactment.

When summarising the effect of authority relations upon WPBA it is important to acknowledge that when an assessment is realised, it will be jointly constructed despite there being distinct differences between trainees' and trainers' approaches. Thus an assessment must be initiated for it to occur, and perhaps such initiation may be seen as the apex of the trainee's power, from which other factors such as case selection will follow.

However, the trainee's power over initiating the episode was not unchallenged, as some interviewees advocated measures such as random case selection, a trainer-led process or

joint leadership of WPBA, with both parties responsible for initiating assessments and choosing cases. There is also the caveat (as noted in the introduction to chapter six) that trainers may initiate unscheduled assessments.

Some trainers clearly wished to lead the process, whilst trainees more often advocated joint leadership. In either case, it appeared that there was a desire on the part of some interviewees to place more power in trainers' hands, to a greater or lesser extent. As well as initiation of the assessments, this also extended to case selection, placing all but the choice of assessor in trainers' hands.

There appeared to be multiple reasons for interviewees desiring this shift of responsibility from trainees to trainers to occur. Some trainees felt it would stimulate trainers' interest in the process, and representatives of both groups mentioned that it would aid planning of assessments and spacing WPBA appropriately during a period of training. Some trainers appeared keen to extend their control of WPBA as a means of countering what they saw as abuse of the process, such as trainees choosing assessors perceived as generous, easy cases or asking for assessments in retrospect.

From the preceding paragraphs, whatever their motives, it is clear that some interviewees advocate greater responsibility for the WPBA process on the part of trainers. As suggested above (and in the literature review) some trainers saw this as a means of reducing perceived abuses of the process by trainees, inevitably accompanied by a greater degree of trainer leadership and ownership of the process.

A potential alternative means of countering the abuses of the process was suggested by Playford et al (2013), who noted that the undergraduates they studied tended to choose

specialists who were known to be hard markers when undertaking WPBA. This was felt to be due to such marking encouraging better performance by the students, within the security of a mentoring relationship. The selection of specialists known to be hard markers contrasts with the opinions expressed by some interviewees in the study, who felt that trainees might select assessors perceived as generous. This might reflect trainees feeling there was a lack of secure mentoring relationships in comparison to the students' situation, or possibly perceiving WPBA in postgraduate medical education as a higher stakes activity and consequently wanting the results recorded in their portfolios to be good.

In terms of the episodes themselves, the interviews suggested that once the assessments have started, there is a definite shift of authority to the assessor, and this is generally retained to the end of the episode. Enquiry regarding trainers wanting trainees to do things their way during WPBA revealed that the majority of trainers were fairly relaxed about this (as long as the trainee was working safely), although a minority expected trainees to follow the assessor's practice, bringing a more overt display of their authority to the assessment.

Thus it would seem that trainers' influence gradually increases to the end of the process, when they decide the outcome of the assessment. This is the direct opposite of the apex of trainees' influence, and it is clear that some trainers would like greater control of both the start and finish of WPBA. Should this occur, the balance of power and control would change, with trainers dominant in all areas of enactment of the assessment, once the trainee had chosen their assessor. However, should trainees retain such choice, they will

still keep an important influence upon the process, as who assesses them has a great influence upon how an assessment is constructed.

The trainee's choice of assessor alludes to a more nuanced view of authority relations during episodes of WPBA. Beyond the steps and rules of the process itself there are likely to be multiple influences upon power and control within it, including the seniority of the trainee (and assessor), prior knowledge of each other, previous shared assessment episodes and the underlying character of both parties. These factors may influence authority relations between individuals during WPBA to a very large degree, but may also be both transient and inconsistent (depending on who is being assessed, the mood of the two participants etc) making it difficult to capture underlying messages.

The main message regarding authority relations between trainer and trainee during episodes of WPBA is that they are subject to multiple influences from the participants themselves, with important consequences for the conduct of the assessments and, potentially, their outcome. Although central guidelines for the conduct of WPBA may appear to confer control to either party at different stages of the process, these are subject to adaptation and challenge by some users, with some trainers wanting greater control. However, how an assessment is constructed (and who controls it) depends on both parties, with numerous human factors outside the guidelines playing their part in deciding where authority lies during assessment, and how much devolves to the trainee.

d. Gaming and subversion in WPBA

The use of the terms gaming and subversion with regard to WPBA may appear to overstate this aspect of the process, but in the data analysis there are numerous examples of how users may manipulate the assessment system to try and meet their own goals and purposes, and this section's title might be defined as such manipulation. Aside from the influences and frank adaptation already discussed, users' goals and purposes occupy a very wide spectrum, from alleged non engagement (on the part of trainers) due to time pressure and high workload, to undertaking numerous, peremptory assessments (on the part of trainees) with the aim of accumulating sufficient to meet an annual target.

This section commences with a discussion of users' engagement in the process, including those who seek to remain unengaged (trainers) and those who pursue limited engagement (trainees), just sufficient to meet annual targets. Discussion of other means of manipulating the system (utilised by trainees, trainers, and both groups) follows, with analysis of the underlying reasons for users' behaviour.

Although those who would not participate in the assessments were not represented, some trainers in the study cited mildly negative feelings regarding the introduction of WPBA, commenting with an air of resignation rather than resistance. Though they participated in the WPBA scheme, it serves to illustrate that there was a wide range of trainer engagement, including (allegedly) nil, limited and enthusiastic involvement.

A narrower range of trainee engagement is unsurprising, as WPBA is a compulsory part of training and thus non participation was not an option. Within the group interviewed,

engagement ranged from successful senior trainees doing enough to get by, to those who were very enthusiastic about the scheme, and clearly exceeded the minimum number of assessments. It seems likely that if the element of compulsion was removed, the range of trainee engagement would be similar to that of the trainers, particularly if it had been possible to interview trainees who were disengaged or performing poorly.

In addition, some interviewees mentioned trainees undertaking peremptory assessments in order to build numbers quickly, suggesting that the length and/or quality of assessment episodes might be minimised in order to promptly achieve a numeric target. Although radiology trainees have (on average) completed more than the minimum target numbers of assessments (Booth, 2012), the quality of these episodes is unknown and compulsion may have played some part in the achievement of these figures.

Data regarding engagement from other specialties is not encouraging, with reports of trainee experience of WPBA in both surgery (Pereira and Dean, 2009) and psychiatry (Menon et al, 2009) suggesting that there were problems in both specialties. WPBA is compulsory in both and 41.4% of surgical respondents reported that the time taken to record results had impacted negatively on their training, whilst the majority of psychiatric respondents felt WPBA had no benefit to their practice, supervision, training or confidence.

In addition a survey of anaesthetic trainers and trainees undertaken by Bindal et al (2013) found that direct observation of procedural skills (a commonly used WPBA tool) was not viewed as a useful training method or learning opportunity. Those anaesthetists surveyed felt the assessments represented a tick-box exercise whose results did not truly

reflect how well a trainee could perform a given task. When one considers the views some trainees hold regarding the value of WPBA (as quoted above), it is perhaps not surprising that the assessments are subject to gaming and subversion.

The extent to which trainers and trainees engage with the process only represents one facet of what might be termed as gaming, and there were various other behaviours allegedly brought to bear on WPBA by trainees, including selecting assessors perceived as generous, choosing unchallenging assessment cases, asking for assessments in retrospect, and undertaking assessments at the end of attachments, when they were likely to perform well.

The motives for the behaviours described in the previous paragraph appear separate to meeting numeric targets, but rather concern passing assessments and achieving positive rather than useful feedback. As early as the first trial interview, it was pointed out that most doctors are orientated towards success in examinations and very averse to any hint of failure, explaining the desire of many trainees to pass the assessments, despite their formative nature being acknowledged by all trainee interviewees. All of the behaviours noted in the previous paragraph concern achieving a good (rather than useful) result, and as it is the result which is recorded in a trainee's portfolio after an individual assessment; it is perhaps not surprising that trainees seek positive feedback.

Although the methods discussed above represent means by which a trainee might achieve a good assessment, a more subtle method mentioned was adapting to the trainer's own perceived likes and dislikes. This could be seen as another facet of manipulating the assessment process, as aside from aiming to perform correctly,

trainees' motives in undertaking an assessed episode the trainer's way are likely to include minimising any possible discord from adopting a different method and demonstrating that they have learnt the assessor's own technique.

By so doing, the trainee might hope to play their part in constructing a harmonious episode contributing to their receiving positive feedback, and such behaviour is not confined to radiology, as demonstrated by Al-Kadri et al (2013), who found that medical undergraduates at their institution tended to map their studies to their supervisors' personalities, presumably for similar motives.

Although trainees' and trainers' differing perspectives are reflected in the ways they may seek to manipulate or amend the assessment process, there are instances where the two groups share common interests and others where they sometimes collude to achieve an outcome desired by both parties. An example of this is when trainers willingly provide a peremptory assessment to trainees looking to increase WPBA numbers prior to their annual review.

By contrast, an example of trainers' and trainees' differing perspectives was the trainers' utilisation of WPBA as a more judgemental process, evidenced by some being prepared to use it to identify failing trainees. This was discussed in greater detail in section a of this chapter, but in common with other adaptations to the process discussed in section b, serves to illustrate that trainers may also adopt behaviours which serve their own purposes.

Although both trainees' and trainers' utilisation of WPBA is influenced by external influences such as central guidelines and the need to meet numeric targets, it appears

that representatives of both groups have other purposes in mind, such as some trainees' desire to pass their assessments. Although some manipulation of WPBA may be due to practical issues such as time constraints, the underlying beliefs of trainers and trainees seemed to have more influence upon how the parties seek to achieve their own ends.

The reasons leading to manipulation of the assessments vary between trainers and trainees. Trainers have more autonomy regarding whether or not they wish to engage in WPBA, and are sometimes able to undertake and adapt assessments to suit their own purposes despite potential conflict with central guidance. By contrast, trainees do not have such autonomy and cannot choose whether to engage or not, but may decide to limit their engagement. Whilst trainees must abide by the rules of WPBA, some seek to use them (e.g. choice of case and assessor) to their own advantage.

It is difficult to compare other professions' manipulation of WPBA with those of the study group or other trainees in postgraduate medical education, as the differing terms of utilisation, degree of compulsion and purposes for which WPBA is used outside medicine makes comparison difficult. In addition there is a paucity of literature regarding the reasons for individuals' engagement in WPBA outside medicine.

However, it is possible to identify in both Poulos and Mahony's (2008) study of undergraduate students and Sambell and Hubbard's (2004) of non traditional higher education students that both first year undergraduates and new non traditional students (respectively) engage in formative assessment hoping for emotional support and help in integrating into university.

As with the examples cited from radiology, this suggests that both student groups have ideas of what they wish to gain from formative assessment, rather than just passive acceptance of the feedback they are given. Although it is not stated whether or not the students' hopes were fulfilled, the desire for assessment to meet specific needs (as in radiology) might encourage users to manipulate the process to try and meet those goals.

e. Summary

Through the majority of this chapter, and in particular the sections concerning WPBA's purpose, its adaptation, power and control during the process, and subversion of the assessments, the unifying message is that central guidelines are widely adapted before WPBA is undertaken in practice. This occurs to varying degrees, and a secondary theme is the amount of adaptation which occurs, ranging from very little to a great deal.

There may be good reason for such adaptation, such as gaps in central guidance or unalterable local circumstances, but the result is that WPBA is a devolved process which cannot be completely uniform when it is enacted between trainers and trainees. This contrasts with assessments such as professional examinations, which are realised centrally and subject to much greater control, leading to a process subject to far less variation.

Searching the literature in order to compare assessment practice in other domains with that in postgraduate medical education led to the realisation that WPBA in medicine tends to dominate what is written on the subject. Some material concerning other areas is available, and sometimes provides an interesting contrast to the situation in

postgraduate medical training by demonstrating the wide range of usage of WPBA in other areas, from completely formative to completely summative, depending on the context in which it is used.

There is little written about end users' adaptation of WPBA from central guidelines and the adaptation literature tends to concentrate upon other forms of policy, where the particular significance of end users to policy implementation is highlighted and amendment of central policy prior to its realisation locally is viewed as almost inevitable in some domains.

With this background, it is probably naïve to think that the RCR did not expect their WPBA policy to be adapted in various ways, and thus allowed for a degree of interpretation of central guidelines. However the experiences from other areas serve to emphasise the importance of engaging with, supporting and training end users in order to realise the main tenets of central policy. This is particularly important for fundamental issues such as the formative purpose of WPBA.

In contrast to the data obtained during the current study, that published from other specialties (psychiatry, surgery and anaesthetics) were predominantly negative. In particular, psychiatry trainees raised fundamental concerns regarding WPBA, perceiving it as; '*conceptually flawed, based on scant evidence and of dubious validity as an assessment tool*' (Menon et al, 2009), and only 11% of those surveyed felt WPBA was backed by good evidence.

The results of these surveys and the generic issues raised (many in common with radiology) served as a means of obtaining a wider picture of trainees' views of WPBA

whilst helping to identify issues which rang true with respondents' views of the radiology scheme. Thus by including views from four specialties (radiology, psychiatry, anaesthetics and surgery) in the discussion, it might be hoped that recommendations for improving generic aspects of WPBA may have greater generalisability than if they were all sourced from radiology alone. These generic issues are outlined in the concluding chapter.

10. Strengths and limitations of the study method

This chapter commences by reviewing how the interview process occurred in practice, and outlines some of the occurrences or influences which I felt contributed to successful data collection. Interview stems which proved less successful are exemplified, followed by examples of unexpected areas which came up during the process, the means I used to probe them, and how I felt I might have improved this. This is followed by a comparison of the study methodology with those of surveys of WPBA undertaken in other medical specialties, and how the method used may have influenced the results. The chapter concludes by discussing what I might have done differently, were I to repeat the study.

When undertaking the interviews, I did not notice any difference in maintaining dialogue between the trainer and trainee groups, and all of the encounters felt fairly relaxed. However, one could sometimes tell that there was a mild degree of tension during some interviews, by observing some interviewees visibly relaxing once recording stopped. I suspect the mild tension during the interviews was due to interviewees feeling they were part of a formal, recorded process and they were a little relieved when it ended. Though interviewees sometimes added remarks afterwards, I did not write down any unrecorded material or include it in the analysis, as this would have felt like obtaining data by subterfuge.

Most of the longer interviews (over an hour) took place later in the study, although this was not universal, and the last interview was one of the shorter ones. Although increasing numbers of issues and confidence on my part in pursuing them with

interviewees may have been a contributing factor toward longer encounters, the length of each interview was determined by many factors, including the loquaciousness of the interviewee and the rapport I was able to establish with them. Establishment of a rapport represented a critical factor in whether or not I could address issues in depth, and was more important than the length of the interview in determining the breadth and depth of the information I could obtain from the encounter. I felt that I was generally able to achieve such a rapport with the interviewees.

Although the interview content evolved as the study progressed, they were all based upon the original schedule and a consistent process was maintained. Some stems, such as asking about the effect of WPBA upon professional relationships tended to elicit little or no response, and information on this topic tended to be obtained indirectly, from answers to other questions.

Examples of this occurred in two trainer interviews, where in each case the interviewee said they felt that WPBA had not affected professional relationships in their departments. Yet both interviewees consistently referred to the trainees with whom they undertook WPBAs as 'candidates', suggesting a more formalised and exam orientated relationship with trainees when they were undertaking assessments than might normally be the case.

Inevitably, as the study developed, areas of interest tended to arise which had not been anticipated when I wrote the interview schedule. Particular examples were who initiated the assessment process and whether cases were chosen by the trainer or trainee. Both

areas were important to interviewees, and contributed information about their attitudes to WPBA as well as how they used the process.

Rather than change the interview schedule to accommodate these new areas, I tended to steer interviewees towards these subjects from my pre existing interview stems in order to preserve the consistency of the interview schedule. In hindsight, after I had read all the transcripts, it might have been better to adapt the schedule as the study proceeded, as there were areas I wish I had pursued with more interviewees and in greater depth, an example being the question of whether or not trainees made a special effort for assessments rather than just having their normal practice observed.

One might derive a view as to whether or not trainees made a special effort for the assessments from their answers to other questions, but I wish I had addressed the subject more directly, as I think that issues such as trainees making a special effort would have a direct bearing upon how they engaged with and enacted the assessment process.

It is possible that a completely anonymous means of data collection regarding the use of WPBA in radiology may have produced slightly different results, as questionnaire surveys undertaken in other specialties such as psychiatry and surgery (outlined in the discussion) resulted in more negative views of the process. Though it is possible respondents might have been more prepared to offer more negative views regarding WPBA in radiology had I undertaken a questionnaire survey rather than interviewing them, there were potential disadvantages to this approach such as needing to define issues from the start, a potentially poor response rate and reduced opportunities to pursue issues in depth compared to semi structured interviews.

Though the preceding paragraphs allude to some potential weaknesses of the chosen method, even with the benefit of hindsight, I would not have conducted the study by an alternative means, particularly because of the opportunity to pursue issues in depth whilst interviewing.

It might have been possible to use other approaches to triangulate the interview data, in order to evaluate WPBA by different methods, and hopefully increase the validity of the study by demonstrating similar findings reached by different means. However, I decided against undertaking this, principally as those willing to join the study comprised a fairly small pool, and I felt that using other means of data collection might adversely affect both interviewees' availability for, and the content of, subsequent interviews.

Thus if I were to repeat the study, I would not alter the primary method, but the two main improvements I would hope to institute would be to obtain responses from those who were disengaged from (or unenthusiastic about) WPBA and to amend the interview schedule to address issues raised during the study which I hadn't anticipated at the start. Recruiting those who were unenthusiastic about WPBA to join the study may still have been challenging, but use of a reminder letter might have helped, and even one such respondent could have provided an important alternative perspective upon the assessment process.

11. Conclusions

This study aimed to address three research questions, and the focus of the first section of this chapter is to reflect upon the research questions, and to outline the main findings from the study which help to address them. As well as reflecting upon the questions, similarities and contrasts between trainers' and trainees' responses are also included where relevant. The second section concerns the implications that the study findings may have for WPBA in higher professional training, recommendations upon how the assessments might be improved in radiology and suggestions for future research in the area.

a. Reflecting upon the research questions

1. How has the policy of introducing Workplace Based Assessment into clinical radiology been realised in practice?

Users generally realised the assessments as formative episodes, even if (as in the opinion of some trainers) they believed that WPBA could have a summative function, particularly as a collection of assessments. As per central guidance, trainers generally assessed episodes they had observed in their entirety and delivered feedback promptly, although trainee interviewees were able to cite instances in which only part of their work was observed or feedback was delayed.

The actions of either trainers or trainees could lead to an episode viewed as unsatisfactory by the other party, with some assessors suggesting that some trainees

engaged with the assessments primarily to build numbers, and did not reflect upon the episodes or incorporate them into a process of long term learning. Conversely, trainees cited examples of trainers completing assessments in a hurried fashion, or in one instance, alleged unfair treatment during an assessment (quoted in discussion).

Although instances of episodes of WPBA being adversely affected by the actions of one or other party to the assessment were cited, reports of disagreement or disharmony between trainer and trainee affecting how an episode was enacted were rare. Such events might be under reported to an interviewer, but there appeared to be little overt conflict at the time of the assessments which might have affected how they were realised.

At the other end of the scale, users quoted examples of what they viewed as good practice, where both trainer and trainee committed to the formative ethos of WPBA, gave sufficient time to undertake and discuss the case, followed by the provision of prompt, specific feedback which was valued and utilised by the trainee. It is likely that the enactment of many assessments fall between these extremes, with the degree of engagement of the two parties and how they interact pushing the episode toward either a satisfactory or unsatisfactory outcome.

Although WPBA was designed as a trainee-led process and there was evidence that many assessments were enacted that way, others were initiated jointly or even at the behest of a trainer. The latter is explicitly permitted by the curriculum and appeared to represent some users' preferred method when realising assessments.

Although the scheme's central tenets were generally preserved, there was widespread local adaptation of the process, and this appears to represent the way in which WPBA

has been realised in radiology practice. The influences (including frank amendment) users exert upon the enactment of WPBA are addressed by the second research question, which follows.

2. How do the different participants engage with, and influence the enactment of, the Workplace Based Assessment process?

There are varying degrees of engagement with WPBA, with a wider range in trainers than trainees, as outlined in the discussion. Trainers and trainees have a very large influence upon the enactment of WPBA, consequent upon how individuals construct assessment episodes between them, how they seek to manipulate it, and how they consciously adapt the process.

Manipulation of the assessment process was also examined in the discussion, with data suggesting that both trainees and trainers undertook this practice for their own purposes. It was suggested that some trainees might select generous assessors, choose easy cases (sometimes chosen in retrospect) and undertake WPBA late in an attachment in order to obtain positive feedback and meet numeric targets.

Conversely trainers might seek not to engage with the process, or undertake peremptory assessments which did little more than increase trainees' numbers. Some trainers brought a more summative edge to the process in terms of assessing competency and identifying failing trainees. Trainers might also consciously amend central policy by using WPBA more summatively depending on trainees' seniority or when they were coming near to the end of a period of training, or not recording the results of poor assessments.

When considering Darling-Hammond's (1990) view, in the context of implementing instructional policy, that adaptation of central guidance is almost inevitable when it is enacted locally, it is perhaps not surprising that WPBA is subject to amendment by end users. This may not be a problem if agreed and seen as fair by those engaging in an assessment, but the study did highlight an instance where a trainee felt unfairly treated. There was also evidence that some trainers viewed WPBA as having a summative purpose in some circumstances, conflicting with the formative ethos of central policy, which was strongly supported by the trainees.

3. How do participants perceive the utility of the system of workplace based assessment which has been developed and how do they think it could be improved?

Judgements of the utility of WPBA are influenced by perceptions of its reliability, validity, feasibility and educational impact. Although interviewees did not always comment in such a specific fashion, some of the paragraphs which follow broadly align with these terms, the first being concerned with the assessments' validity whilst those in which interviewees suggest improvements to the scheme are more allied to improving WPBA's feasibility.

Although there were members of both groups who thought that WPBA could be used to assess complex work in radiology, there were others who took a contrary view and aligned themselves with Grant's (2007) view, that the assessments were only suitable for assessing simpler activities. Even amongst those who thought WPBA suitable for assessing complex activity, there was more enthusiasm for the assessment of technical

or procedural steps, rather than less well defined behaviours such as leadership and adapting to changed circumstances.

Weaknesses of WPBA identified by interviewees included the perception that the assessments represented a tick-box exercise, time pressures making it difficult for assessors to undertake them, failure to identify poorly performing trainees and assessment scoring forms which were too generic in nature.

A further weakness of WPBA identified in both this study and those of other specialties (e.g. psychiatry) was of poor assessor preparedness and training. In an invited commentary upon cited articles regarding the introduction of WPBA into psychiatry by Menon et al (2009) and Babu et al (2009), Oyeboode (2009) found that inadequate training of assessors when the scheme was introduced caused similar issues to those highlighted in this study. It is possible that as the interviews were conducted shortly after WPBA was introduced into radiology, and assessor training is available, the situation may have improved in the interim, but as in psychiatry, there is still likely to be room for improvement.

In terms of how WPBA might be improved, the most popular suggestion was to ensure there was protected time for WPBA, and interestingly, this was only suggested by trainers, suggesting that they might feel time pressures more acutely than trainees.

Marginally less frequently, the improvements most suggested by both groups were that WPBA should be either trainer or jointly led, with a degree of pre planning between trainers and trainees.

Another improvement suggested by interviewees related to the use of more specific assessment forms for individual procedures rather than the generic WPBA forms in current use. This suggestion has resonance with another area mentioned in this section, namely the assessment of complex work, by developing and utilising specific Procedure Based Assessments to assess trainees' performance of certain examinations. These are being trialled for future use, but might improve the assessment of complex activity by allowing specific comment to be made upon the multiple stages of a procedure.

Interviewees from both groups also suggested that joint leadership could be a means of improving the engagement of both parties, with potentially important effects such as ensuring appropriate cases were chosen for assessment, selection of material in prospect, and planning a programme of WPBA in advance, allowing formative assessments to be spread evenly throughout a period of training.

b. Implications of the study for future practice and research

In this section I have attempted to refine the study's most important messages and derive recommendations to improve future WPBA practice in radiology. In addition, published data regarding the utilisation of WPBA in other medical specialties is compared with that from the current study, in an attempt to extract generic messages which may be more generally applied. Finally, I have made suggestions for future research in the field.

The overriding message which has emerged from the study is that central guidelines regarding WPBA may be widely influenced or adapted by users for numerous reasons when the assessments are realised in the workplace. Material regarding implementation

of central policy by end users from other domains (as quoted in previous chapters) suggests that such adaptation is almost inevitable and may be important for it to succeed.

Examples of such adaptation applied to WPBA might include allowing a trainee's work to be assessed by a small number of assessors if resources or a department's size are limited or by making allowances for the workplace's own demands upon trainers and trainees by occasionally allowing delayed feedback delivery when they are hard pressed by service commitments.

This study has shown that there is variation in day to day practice by individuals when WPBA is realised, and although the central tenets of the process are respected in most cases, there are a minority in which practice may lie outside them. The boundaries of a normal range of practice are difficult to define exactly, but central guidance might be regarded as a start point.

If such adaptation is inevitable, it is important to try and differentiate which amendments to central policy are desirable or allowable and which are undesirable, and should be discouraged. An example from the study of a desirable change might be instituting joint trainer and trainee leadership of WPBA, whilst undesirable changes might include those which depart from the main tenets of central guidance and/or disadvantage trainees, such as not recording assessments or undertaking them without the trainee's knowledge.

The purpose and importance of WPBA are linked to the question of adapting central guidelines, because if the assessments were regarded as purely formative episodes

whose results were not formally recorded then end users might agree to amend the process to suit the particular needs of an individual assessor and trainee, with very wide variation depending on their circumstances. However, this is not the case in practice, as the results are scored, form part of a trainee's record and may be seen by others. Even if the assessments were not scored as per the supervised learning events (Royal College of Radiologists, undated) mentioned later in this section, and the results expressed as a narrative, they would still be recorded.

The scoring of episodes and recording of the results of the assessments inevitably raises the stakes in WPBA, making adaptation by individuals outside the normal range of the process difficult to allow. This raises the question of what normal range is allowable, when the resources, facilities and staffing of training institutions will vary widely. It is likely that some allowance should be made for such variation, but the main tenets of central guidance (see next paragraph) should be respected and any actions or amendments which disadvantage trainees should lie outside the normal range.

I think it is very likely that the RCR was cognisant of such variation when it issued its guidelines, as a national process would have to be deliverable by all training schemes. However, there are some central tenets of WPBA which they might regard as non negotiable, such as its formative ethos, the types of assessments to be performed, how the results are recorded and the numeric targets to be met. Aside from some trainers' desire for some summative usage of WPBA, these principles were generally not challenged by the interviewees, and I would concur with their importance, based upon the requirement for a formative process whose results form part of a trainee's record.

Thus provided that the assessments are used in a formative, low-stakes fashion, and are seen as fair by both parties involved, central guidance on the main principles of WPBA with the allowance of some discretion in their detailed implementation may be the best means of allowing the process to be best realised under differing local circumstances. The advantages of this balanced approach are summarised by Gunderman (2012), who suggests that in the realm of educational assessment;

‘Centralization promotes fairness and efficiency, while decentralization promotes engagement and creativity. Letting the pendulum swing too far in either direction is fraught with peril’.

Local adaptation beyond the norms suggested above could be addressed when training users (both trainers and trainees) by emphasising the RCR’s view of the formative purpose of WPBA, whilst aiming to optimise the conduct of the assessments and delivery of feedback. Such training might also help reduce the peremptory, tick-box nature of some assessments, though whilst numeric targets remain in place, they are unlikely to disappear altogether.

The peremptory nature of some assessments could also be addressed by redesigning the score sheets to remove elements of ranking or grading. This is set to occur in the Foundation Programme with the introduction of completely formative Supervised Learning Events which cannot be scored, ranked, passed or failed (Kessel et al, 2012).

In addition to ensuring trainers (and trainees) were trained in WPBA, from my own experience, I would advocate widening available trained faculty to include senior trainees and allied health professionals (e.g. radiographers) in order to try and spread the

assessment load and relieve some pressure on those assessors who feel they are time poor. The use of such wider faculty is specifically authorised in central guidance and already occurs (to some extent) locally.

In addition to optimising training and widening assessor faculty, from this work it appears that the most practical step which could be taken to improve the day to day realisation of WPBA in radiology would be to adopt joint leadership of the process (trainer and trainee) and have a pre planned programme of assessments during an attachment agreed by both parties. These measures would hopefully reduce opportunities for the manipulation of WPBA and ensure that the assessments were properly planned and spaced through a period of training.

In terms of aligning improvements to WPBA recommended by this study with postgraduate medical training as a whole, there is literature based upon the assessments' use in other specialties such as surgery, psychiatry and anaesthetics. Although they are different specialties to radiology, Sugand et al's (2011) review of WPBA's utilisation in psychiatry, and Bindal et al's (2013) survey of anaesthetists demonstrate some common themes in terms of improving the assessments.

As was suggested by the current study, Sugand et al (2011) recommend that assessments are planned at the start of clinical attachments, spaced out through them, and trainers given protected time in which to undertake WPBA. Bindal et al (2013) recommend that both trainers and trainees are trained in use of the assessments and concurred regarding the need for planning and allocating time for WPBA. These may be viewed as training issues which are common to most specialties, and although there was less common

ground between specialty specific practices, referral to chapter 8 suggests that the issues regarding WPBA raised by interviewees were often generic rather than confined to radiology.

In terms of further research based upon the findings of this study, if it were possible to action some of the recommendations made in this section (e.g. joint leadership of WPBA, greater training of assessors, wider faculty) it would be useful to interview trainers and trainees in future to see if it improved their perception of the assessment process.

If central guidelines were strengthened or made more prescriptive in future, perhaps by mandating assessment of certain types of cases and/or their complexity, it would be interesting to assess the degree of adaptation of WPBA which end users might still pursue. Were the central guidelines loosened, perhaps by removing scoring and/or not recording the results, it would be interesting to see if end users could fully embrace the formative nature of the assessments and whether or not manipulation of WPBA would still occur.

It is important to realise that this study is based upon the stated views of interviewees, and although positive opinions were expressed regarding some aspects of WPBA, it is uncertain whether it leads to improved performance. Similarly, within the wider literature, it has not yet been possible to prove that these assessments have led to an improvement in doctors' performance in terms of attitudes, skills, knowledge or behaviour (Miller and Archer, 2010). As WPBA has been universally introduced into

postgraduate medical training, it would seem vital that future research beyond radiology is undertaken in order to prove or disprove its efficacy.

It may take a very long time to demonstrate whether or not those practitioners who have been assessed with WPBA during their training have had their performance improved by using the process and are able to maintain accepted standards when they practice independently. Additionally, there will inevitably be numerous other factors in play during training whose influences upon performance will be very hard to separate from those of WPBA.

It may be possible to control for these additional factors by comparing the outcomes of training between schemes which use WPBA to a greater or lesser extent, but are otherwise similar. One might not expect greatly differing extents of WPBA utilisation between training schemes when the assessments are a country wide initiative, but such variation does occur, and it is possible to identify differing usage between centres.

It should also be noted that as WPBA was only introduced into radiology in 2010, no trainee in the specialty has yet completed their five year training under its auspices from start to finish. Once trainees have undergone their entire training subject to WPBA, it may prove easier to derive messages regarding its long term efficacy rather than by utilising the preliminary data relating to radiology which is currently available.

In my view, although it is not yet possible to prove the efficacy of WPBA in radiological training, the assessment system should be pursued, to both fill the gap in the assessment of trainees' performance which previously existed (and for which no other

solution has been proposed) and to build the evidence base to support the utilisation of the process in the future.

12. References

1. Academy of Medical Royal Colleges. (2009) Improving Assessment. Academy of Medical Royal Colleges, London.
2. Al-Kadri H M, Al-Kadi M T and Van der Vleuten C P M. (2013) Workplace-based assessment and students' approaches to learning: a qualitative inquiry. *Medical Teacher*; 35(s1): s31 – s38.
3. Augustine K, McCoubrie P, Wilkinson J R and McKnight L. (2010) Workplace-based assessment in radiology – where to now? *Clinical Radiology*; 65: 325-332.
4. Babu K S, Htike M M and Cleak V E. (2009) Workplace-based assessments in Wessex: the first 6 months. *Psychiatric Bulletin*; 33: 474-478.
5. Bindal N, Goodyear H, Bindal T, and Wall D. (2013) DOPS assessment: A study to evaluate the experience and opinions of trainees and assessors. *Medical Teacher*; 35(6): e1230-e1234.
6. Booth J. (2012) WPBA Use Clinical Radiology Aug 2011-July 2012. The Royal College of Radiologists, London (unpublished, internal document).
7. Boud D. (2000) Sustainable Assessment: Rethinking assessment for the learning society. *Studies in Continuing Education*; 22(2): 151-167.
8. British Educational Research Association (2004) Revised Ethical Guidelines for Educational Research. BERA (Chair J. Gardner) Southwell, Notts.
9. Brodie P and Irving K. (2007) Assessment in work-based learning: investigating a pedagogical approach to enhance student learning. *Assessment & Evaluation in Higher Education*; 32(1): 11-19.

10. Burkill G J C. (2008) Work-based assessment for trainees – more than just a few new tools? *Clinical Radiology*; 63: 12-14.
11. Chrispeels J H. (1997) Educational Policy Implementation in a Shifting Political Climate: The California Experience. *American Educational Research Journal*; 34(3): 453-481.
12. Cohen D K and Ball D L. (1990) Policy and Practice: An Overview. *Educational Evaluation and Policy Analysis*; 12(3): 233-239.
13. Cohen L, Manion L and Morrison K. (2000) *Research Methods in Education*. RoutledgeFalmer. London and New York.
14. Costley C. (2007) Work-based learning: assessment and evaluation in higher education. *Assessment and Evaluation in Higher Education*; 32(1): 1-9.
15. Crossley J, Johnson G, Booth J and Wade W. (2011) Good questions, good answers: construct alignment improves the performance of workplace-based assessment scales. *Medical Education*; 45: 560-569.
16. Darling-Hammond L (1990) Instructional Policy Into Practice: “The Power of the Bottom Over the Top”. *Educational Evaluation and Policy Analysis*; 12(3): 339-347.
17. Delandshere G and Petrosky A R. (1998) Assessment of Complex Performances: Limitations of Key Measurement Assumptions. *Educational Researcher*; 27(2): 14-24.
18. Denison A. (2010) Workplace Based Assessment. Report of the National Pilot 2008-2009. The Royal College of Radiologists, London. Available at: www.rcr.ac.uk/docs/radiology/pdf/WpBAPilotSum-FINAL.pdf

(Accessed 14/06/10)

19. DiCicco-Bloom B and Crabtree B F. (2006) The qualitative research interview. *Medical Education*; 40: 314-321.
20. Dunsire A. (1990) Implementation Theory and Bureaucracy. In Younis T (Ed) *Implementation in Public Policy*. Dartmouth. Aldershot, Brookfield USA, Hong Kong, Singapore, Sydney.
21. Eraut M. (2004) Informal learning in the workplace. *Studies in Continuing Education*; 26(2): 247-273.
22. Firestone W A, Mayrowetz D and Fairman J. (1998) Performance-Based Assessment and Instructional Change: The Effects of Testing in Maine and Maryland. *Educational Evaluation and Policy Analysis*; 20(2) 95-113.
23. Fox G K, Bowden S C, and Smith D S. (1998) On-Road Assessment of Driving Competence after Brain Impairment: Review of Current Practice and Recommendations for a Standardized Examination. *Archives of Physical Medicine and Rehabilitation*; 79(10): 1288-1296.
24. Francis I S. (2008) Finding and maintaining professionalism in radiology. *Clinical Radiology*; 63: 15-17.
25. Fullerton H. (2003) Observation of teaching. In Fry H, Ketteridge S and Marshall S (Eds). *A Handbook for Teaching and Learning in Higher Education*. RoutledgeFalmer. London and New York.
26. Gaba D M, Howard S K, Flanagan B, Smith B E, Fish K J and Botney R. (1998) Assessment of Clinical Performance during Simulated Crises Using Both Technical and Behavioral Ratings. *Anesthesiology*; 89(1): 8-18.

27. General Medical Council (2010) Workplace Based Assessment: A guide for implementation. General Medical Council, London.
28. General Medical Council (2010) GMC Education Strategy 2011-2013. General Medical Council, London.
29. Gingerich A, Regehr G and Eva K W. (2011) Rater-Based Assessments as Social Judgments: Rethinking the Etiology of Rater Errors. *Academic Medicine*; 86(10): s1-s6.
30. Gonczi A (1994) Competency Based Assessment in the Professions in Australia. *Assessment in Education*; 1(1): 27-44.
31. Govaerts M J B, Van der Vleuten C P M, Schuwirth L W T and Muijtjens A M M. (2007) Broadening Perspectives on Clinical Performance Assessment: Rethinking the Nature of In-training Assessment. *Advances in Health Sciences Education*; 12: 239-260.
32. Govaerts M J B, Van de Wiel M W J, Schuwirth L W T, Van der Vleuten C P M and Muijtjens A M M. (2013) Workplace-based assessment: raters' performance theories and constructs. *Advances in Health Sciences Education*; 18: 375-396.
33. Govaerts M and Van der Vleuten C P M. (2013) Validity in work-based assessment: expanding our horizons. *Medical Education*; 47: 1164-1174.
34. Grant J, Maxted M, Chambers K, and Marshall J. (2007) Workplace based assessment in medicine: Significance within the practice and learning environment and generalisability to other professional practice-based learning contexts: A short project summary. The Open University Centre for Education in Medicine. Available at:

www.open.ac.uk/pbpl/activities/details/detail.php?itemId=460d3238ea7c3&theMeId=460299b584ed5 (Accessed 29/04/2008).

35. Gunderman R B. (2012) Philosophical Considerations in Educational Assessment. In: Hibbert K M, Chhem R K, Van Deven T and Wang S-C (Eds). Radiology Education: The Evaluation and Assessment of Clinical Competence. Springer, Berlin, Heidelberg.
36. Hamman W R (2004) The complexity of team training: what we have learned from aviation and its applications to medicine. *Quality and Safety in Health Care*; 13 (Supplement 1): i72-i79.
37. Harding J and McCoubrie P (2009) Competency-based training versus traditional experience in radiology; how best to educate the radiologists of the future? *Clinical Radiology*; 64: 569-573.
38. Hattie J and Timperley H (2007) The Power of Feedback. *Review of Educational Research*; 77(1): 81-112.
39. Hauer K E and Kogan J R (2012) Realising the potential value of feedback. *Medical Education*; 46: 140-142.
40. Hunt M R, Chan L S and Mehta A (2011) Transitioning from Clinical to Qualitative Research Interviewing. *International Journal of Qualitative Methods*; 10(3): 191-201.
41. Kessel D, Jenkins J and Neville E (2012) Workplace based assessments are no more. *BMJ Careers* 29/9/12: 5.

42. Kogan J R, Conforti L N, Bernabeo E C, Durning S J, Hauer K E and Holmboe E S. (2012) Faculty staff perceptions of feedback to residents after direct observation of clinical skills. *Medical Education*; 46: 201-215.
43. Long G. (2001) Documentation of In-training Assessment for Radiology Trainees. *Clinical Radiology*; 56: 310-315.
44. McLaughlin M W. (1987) Learning From Experience: Lessons From Policy Implementation. *Educational Evaluation and Policy Analysis*; 9(2): 171-178.
45. Menon S, Winston M and Sullivan G. (2009) Workplace-based assessment: survey of psychiatric trainees in Wales. *Psychiatric Bulletin*; 33: 468-474.
46. Miller A and Archer J (2010) Impact of workplace based assessment on doctors' education and performance: a systematic review. *British Medical Journal*; 341: c5064.
47. Miller G E. (1990) The assessment of clinical skills/competence/performance. *Academic Medicine*; 65(9): s63-s67.
48. Moorthy K, Munz Y, Sarker S K, and Darzi A. (2003) Objective assessment of technical skills in surgery. *British Medical Journal*; 327: 1032-1037.
49. National Research Ethics Service (2009) Information Sheets & Consent Forms. Guidance for Researchers and Reviewers. Version 3.5. National Research Ethics Service, London.
50. Norcini J J. (2007) Workplace-based assessment in clinical training. Association for the Study of Medical Education. Edinburgh.
51. Norcini J. and Burch V. (2007) Workplace-based assessment as an educational tool: AMEE Guide No.31. *Medical Teacher*; 29: 855-871.

52. O'Connor N, Kotze B, Vine R, Patton M and Newton R. (2012) The emperor's edicts stop at the village gate. *Australasian Psychiatry*; 20 (1): 28-34.
53. Oyeboode F. (2009) Competence or excellence? Invited commentary on... Workplace-based assessments in Wessex and Wales. *Psychiatric Bulletin*; 33: 478-479.
54. Pell G, Horner M S and Roberts T E. (2008) Assessor training: its effects on criterion-based assessment in a medical context. *International Journal of Research and Method in Education*; 31(2): 143-154.
55. Pereira E A C and Dean B J F. (2009) British surgeons' experiences of mandatory online workplace-based assessment. *Journal of the Royal Society of Medicine*; 102: 287-293.
56. Playford D, Kirke A, Maley M and Worthington R. (2013) Longitudinal assessment in an undergraduate longitudinal integrated clerkship: The mini Clinical Evaluation Exercise (mCEX) profile. *Medical Teacher*; 35(8): e1416-e1421.
57. Poikela E. (2004) Developing criteria for knowing and learning at work: towards context-based assessment. *The Journal of Workplace Learning*; 16(5): 267-274.
58. Postgraduate Medical Education and Training Board (2008) Standards for curricula and assessment systems. Postgraduate Medical Education and Training Board, London.
59. Poulos A and Mahony M J. (2008) Effectiveness of feedback: the students' perspective. *Assessment and Evaluation in Higher Education*; 33(2): 143-154.

60. Ramsden W H. (2009) The Radiology Integrated Training Initiative (R-ITI).
Medical Education; 43(5): 494-495.
61. Rethans J-J, Norcini J J, Baron-Maldonado M, Blackmore D, Jolly B C, LaDuca T, Lew S, Page G G, and Southgate L H. (2002) The relationship between competence and performance: implications for assessing practice performance.
Medical Education; 36: 901-909.
62. Rosenblatt E, Haffty B G and Leer J W (2012) Assessing Clinical Competence in Radiation Oncology Education. In: Hibbert K M, Chhem R K, Van Deven T and Wang S-C (Eds) Radiology Education: The Evaluation and Assessment of Clinical Competence. Springer, Berlin, Heidelberg.
63. Rowntree D. (1987) Assessing Students: How shall we know them?
RoutledgeFalmer, London and New York.
64. The Royal College of Radiologists (Faculty of Clinical Radiology). (2010) Specialty Training Curriculum for Clinical Radiology. The Royal College of Radiologists. London.
65. The Royal College of Radiologists (2010) mini-IPX Guidance for Assessors.
The Royal College of Radiologists, London. Available at:
www.rcr.ac.uk/docs/radiology/other/mini-IPX%20Guidance%20for%20Assessors%20v5.0pdf (accessed 11/01/2013)
66. The Royal College of Radiologists (2010) Rad-DOPS Guidance for Assessors.
The Royal College of Radiologists, London. Available at:
www.rcr.ac.uk/docs/radiology/other/Rad-DOPS%20Guidance%20for%20Assessors%20v5.0pdf (accessed 11/01/2013)

67. The Royal College of Radiologists (2010) mini-IPX and Rad-DOPS forms. The Royal College of Radiologists, London. Available to download at: www.rcr.ac.uk/content.aspx?pageID=1803 (accessed 11/01/2013)
68. The Royal College of Radiologists (undated) Glossary of Educational Terms. The Royal College of Radiologists, London. Available to download at: www.rcr.ac.uk/docs/radiology/pdf/glossary_of_educational_terms.pdf (accessed 01/09/2013)
69. Sadler D R. (2009) Indeterminacy in the use of preset criteria for assessment and grading. *Assessment & Evaluation in Higher Education*; 34(2): 159-179.
70. Sadler D R. (2010) Fidelity as a precondition for integrity in grading academic achievement. *Assessment & Evaluation in Higher Education*; 35(6): 727-743.
71. Sambell K and Hubbard A. (2004) The Role of Formative 'Low-stakes' Assessment in Supporting Non-traditional Students' Retention and Progression in Higher Education: Student Perspectives. *Widening Participation and Lifelong Learning*; 6(2): 25-36.
72. Setna Z, Jha V, Boursicot K A M and Roberts T E. (2010) Evaluating the utility of workplace-based assessment tools for specialty training. *Best Practice & Research Clinical Obstetrics and Gynaecology*; 24: 767-782.
73. Southgate L, Grant J. (2004) Principles for an assessment system for postgraduate medical training (a working paper). Postgraduate Medical Education and Training Board, London.

74. Stanley M, Brooker R and Gilbert R. (2002) Examiner Perceptions of Using Criteria in Music Performance Assessment. *Research Studies in Music Education*; 18(1): 46-56.
75. Sugand K, Palod S, Olua K, Saha S, Naeem A, Matin S and Howlett M. (2011) Systematic review of workplace-based assessments in psychiatry: surgical dissection and recommendations for improvement. *The Psychiatrist*; 35: 419-424.
76. Swanson D B, Norman G R, and Linn R L. (1995) Performance-Based Assessment: Lessons From the Health Professions. *Educational Researcher*; 24(5): 5-11, 35.
77. Wang S-C. (2012) Assessment and Evaluation in a Transnational Radiodiagnosis Training Program. In: Hibbert K M, Chhem R K, Van Deven T and Wang S-C (Eds) *Radiology Education: The Evaluation and Assessment of Clinical Competence*. Springer, Berlin, Heidelberg.
78. Wilkinson J R, Crossley J G M, Wragg A, Mills P, Cowan G and Wade W. (2008) Implementing workplace-based assessment across the medical specialties in the United Kingdom. *Medical Education*; 42: 364-373.
79. Yorke M. (2009) Grading student achievement in Higher Education: Measuring or Judging? In Tight M, Mok K H, Huisman J and Morpew C C (Eds) *The Routledge International Handbook of Higher Education*. Routledge. New York.
80. Younis T and Davidson I. (1990) The Study of Implementation. In Younis T (Ed) *Implementation in Public Policy*. Dartmouth. Aldershot, Brookfield USA, Hong Kong, Singapore, Sydney.

13. Appendices

- a. Tools used in WPBA in Radiology
- b. Semi-structured interview schedule
- c. Participant Information Sheet
- d. Consent Form
- e. Numbers of WPBAs undertaken by respondents
- f. RCR Assessments (extract from curriculum)
- g. mini-IPX scoring form
- h. mini-IPX Guidance for Assessors
- i. Rad-DOPS scoring form
- j. Rad-DOPS Guidance for Assessors
- k. Data Coding System

a. Tools used in WPBA in Radiology

Workplace based assessment in Radiology is undertaken by the utilisation of five tools, and these are as follows:

1. Mini- Imaging Interpretation Exercise (mini-IPX). This involves the assessor observing the trainee interpreting a series of images or examinations, and discussing their findings with them.
2. Radiology Direct Observation of Procedural Skills (Rad-DOPS). This involves the assessor observing the trainee undertaking a practical procedure.
3. Audit Assessment. This involves the trainee presenting a clinical audit to the assessor and discussing the findings.
4. Teaching Observation. This involves the assessor directly observing a trainee's teaching of others.
5. Multi-Source Feedback (MSF). This involves an annual (anonymised) assessment of a trainee's professional behaviours by a sample of colleagues with whom they work.

b. Semi-structured interview schedule

Thank you very much for coming along to speak to me today. As you know, I should like to ask for your views on workplace based assessment in radiology, from your perspective as either a trainer or trainee. I hope the questions I shall ask will seem clear, but if any of them aren't, please ask me to rephrase them. Although the questions may seem fairly specific, I would be very pleased to hear any view you have regarding workplace based assessment, so please feel free to expand upon the issues as we go along.

1. Could you describe your job or role in radiology?
2. What is your involvement with Workplace Based Assessment?
Have you actually undertaken one/them as assessor or as a trainee?

If so, how did you find it?
3. If you were assessed, was it easy to find someone to assess you?
Did you seek out someone in particular? If so, why?

If you were an assessor, how feasible was it to undertake assessments?
4. If you were assessed, were you clear about the role of the assessment?
(Did it feel like a test, an educational exercise, or both?)
5. During your involvement, which of the WPBA tools have you used?
If you were an assessor, were you trained in their use?
6. How did you feel the process went? Was it smooth or were there problems?
Was it practical and feasible? If not, why was that?
7. Did you assess/were you assessed based upon a set checklist, an overall impression, or as a combination of the two?

Which approach do you favour? Why?
8. Do you feel that the process is reliable? Does it feel subject to the assessor's own likes or dislikes or does it feel free of these?

Do you think the results, either as a mark sheet or oral feedback, truly reflect your/ the trainee's performance?

9. Do you think WPBA is suited to the assessment of complex professional activity? Can it address the multiple steps and facets of complex work? If not, why not? If it is seen as partially successful, what areas can be satisfactorily assessed?
10. Can WPBA address areas such as leadership or being a member of a team?
Can it be used to assess behaviours such as adapting to changed circumstances?
11. Did you give/get feedback when undertaking WPBA?
Did you feel this was an important and useful part of the process?
12. What effect do you think WPBA had upon your/your assessee's future learning?
13. Do you think WPBA can be used both developmentally and finally?
Can the tools be used alone, or other forms of assessment needed?

How does the process integrate with other forms of assessment?

Do they address the same things?
14. Overall, do you think the WPBA tools reflect trainees' ability to undertake complex work? Would you use/trust the results (especially if done by others)?
15. Do you feel that the WPBA scheme has had any good or adverse affects upon professional relationships in the department?
Has it affected departmental working?
16. Are there any especially good or poor features of WPBA?
17. Have you any suggestions for improving the process?

c. Participant Information Sheet

Workplace Based Assessment in Higher Professional Training

I would like to invite you to take part in the above study. Please take time to read the following information carefully to see if you would like to participate. Talk to others about the study if you wish, and please ask me if there is anything which is not clear or if you would like more information. Take time to decide whether or not you wish to take part.

I am undertaking this study as part of my Doctorate in Education at The University of Leeds. It will be based upon interviews with consultant radiologists in Leeds and trainees on the Leeds/Bradford radiology training scheme. It will seek to explore their views of workplace based assessment in radiology.

1. What is the purpose of the study?

Its aim is to look at how consultants and trainees view workplace based assessment in radiology, with particular reference to its utility in assessing complex medical activity.

It will help to inform discussion about the use of such assessments in the future, and hopefully lead to improvements where appropriate.

2. Why have I been invited to take part?

The study group will comprise consultant radiologists from the Leeds Teaching Hospitals and trainees on the Leeds/Bradford radiology training scheme. These groups were selected as they will directly participate (or make use of the results) in workplace based assessment.

3. Do I have to participate?

Participation is entirely voluntary. If you do agree to take part, you will be asked to sign a consent form prior to joining the study and a copy of this will be given to you to keep. Even so, you are free to withdraw at any time without giving a reason, and if you do this, any interview data you have provided will not be used. Either withdrawal or declining to join the study would have no affect upon your current status or future progress.

4. What will happen if I decide to take part?

You will be asked to give your views on workplace based assessment in a single interview with William Ramsden. This will last around 45 minutes and will be arranged at a time and place convenient to you. In order to ensure data collection is as accurate as possible the interviews will be recorded and later transcribed. The transcripts will be anonymised, as will any answers they contain which might identify any other individual.

5. What are the possible disadvantages of taking part?

Giving frank or negative views on assessment in radiology to one of the local organisers of training may cause anxiety in terms of causing offence or trainees harming their future career prospects. However, the study's accuracy depends upon interviewees having the confidence to express themselves honestly, and they may be reassured that a negative view will be received as impartially as any other.

6. What are the possible benefits of taking part?

Although the study may not benefit interviewees personally, the information gained from it will help to inform future improvements in workplace based assessment.

7. What happens when the study stops?

The study is due to run during 2011, and will be written up in 2012. A summary of the main findings will be made available to those participants who would like to receive it.

8. What if there is a problem?

Any concern or complaint you may have about how you have been dealt with during the study will be addressed. In the first instance, please contact the investigator, William Ramsden (0113 392 3791). If this is inappropriate, or he is unable to help, you may contact his academic supervisor, Professor Trudie Roberts (0113 343 1657). Indemnity is shared by The University of Leeds and The Leeds Teaching Hospitals NHS Trust.

9. Will my taking part in this study be kept confidential?

I will follow ethical and legal practice and all the study data will be handled in confidence. The anonymised interview transcripts will be kept in a locked filing cabinet and the investigator will be the only person who will have access to them. No other individual will be told whether or not you have participated in the study. The interview data will be used for this study alone.

10. What will happen to the results of the study?

The research will be written up during 2012 and submitted as the dissertation required for the degree of Doctor of Education. The study data will be retained for up to five years afterwards and may be subsequently presented at a radiological or educational meeting, or submitted to a peer reviewed journal. Should either occur, it would be important to note that the anonymity of interviewees would be preserved.

11. Who has reviewed the study?

The study has been approved by The Leeds East Research Ethics Committee.

Contact details for further information

1. Investigator: Dr. William Ramsden. E-mail: william.ramsden@leedsth.nhs.uk

Telephone number: 0113 392 3791

Address: Clarendon Wing X-Ray Dept,
The General Infirmary at Leeds,
Belmont Grove
Leeds
LS2 9NS

2. Supervisor: Professor Trudie Roberts. E-mail: t.e.roberts@leeds.ac.uk

Telephone number: 0113 343 1657

Address: University of Leeds,
Worsley Building,
Clarendon Way,
Leeds
LS2 9NL

d. Consent Form

Workplace Based Assessment in Higher Professional Training

Name of Researcher: Dr. William Ramsden

Please Initial

1. I confirm that I have read and understand the information sheet
for the above study. I have had the opportunity to consider the information,
ask questions and have had these answered satisfactorily.
2. I understand that my participation is voluntary and I am free to withdraw at
any time without giving any reason. This will not affect my current or future
status.
3. I understand that my interview will be recorded and transcribed. The resultant
data will be securely stored and will not be subject to any unauthorised or
illegitimate access.
4. I understand that although the study may be submitted for publication, I will
not be identified. Interview transcripts will be rendered anonymous and my
own data will remain confidential.
5. I agree to take part in the above study.

.....
Name of Date Signature
Participant

.....
Name of Date Signature
Investigator

When completed, 1 copy for participant, 1 for investigator file.

e. Numbers of WPBAs undertaken by respondents

Respondent	WPBAs undertaken	Gender
Trainer 1	10 (approx)	M
Trainer 2	20 (approx)	M
Trainer 3	12 – 15	F
Trainer 4	12 – 15	M
Trainer 5	50 (approx)	M
Trainer 6	12 – 14	F
Trainer 7	8 – 10	M
Trainer 8	10 (approx)	F
Trainer 9	8 (approx)	M
Trainer 10	3	F
Trainer 11	20 (approx)	M
Trainer 12	2 – 3	F
Trainee 1	10 (approx)	M
Trainee 2	14 – 16	M
Trainee 3	20 (approx)	F
Trainee 4	10 (approx)	M
Trainee 5	30 (approx)	M
Trainee 6	17 (approx)	F
Trainee 7	16 (approx)	M
Trainee 8	20 (approx)	M

f. RCR Assessments (extract from curriculum)

Purpose of assessment

The assessment system included in this curriculum is intended to

- enhance learning by providing formative assessment, enabling trainees to receive immediate feedback, measure their own performance and identify areas for development;
- drive learning and enhance the training process by making it clear what is required of trainees and motivating them to ensure they receive suitable training and experience;
- provide robust, summative evidence that trainees are meeting the curriculum standards during the training programme;
- ensure trainees are acquiring competencies within the domains of Good Medical Practice;
- assess trainees' actual performance in the workplace;
- ensure that trainees possess the essential underlying knowledge required for Clinical Radiology;
- inform the Annual Review of Competence Progression (ARCP) is the culminating part of the assessment system;
- identify trainees who should be advised to consider changes of career direction.

Assessment methodology

Continuous assessment

Improvement in clinical practice will only happen if regular review leads to constructive feedback. Thus, continuous review and assessment is a fundamental part of clinical radiology training. Radiology trainees are expected to demonstrate improvement and progression during each attachment. It is anticipated that radiology trainees will increasingly reach higher levels of attainments as they progress through their training. It is important that they arrange and undertake assessments in a timely and educationally appropriate manner spread throughout the year.

Arriving at the overall assessment and judgement of the radiology trainee must be based on multiple assessments by many assessors, on multiple occasions. During core training, within a typical three/four month placement, an individual consultant/assessor is unlikely to build up a coherent picture of competences, let alone performance, of an individual trainee. Therefore, the training programme director (TPD) will ensure that there is a local faculty of trainers capable of building a balanced judgement of a trainee's performance supported by the workplace based assessment results. Such an approach will prevent any individual having undue influence regarding a trainee's progression.

Self Assessment

Radiology trainees have a personal responsibility to undertake self assessment an integral part of their professional life. It is good educational practice for this to be stated clearly and discussed fully during induction.

Assessment System and Tools

Radiological practice will be assessed using an integrated package of workplace based assessments and summative examination of knowledge and radiological skills, which will sample across the domains of the curriculum. The assessment methods are fit for purpose and mapped onto the curriculum in an integrated way.

The assessments will generate structured feedback for trainees within core radiological training and level 1/2 training. The assessment tools have been selected on the basis of their fitness for purpose.

Summative Assessment

The First FRCR Examination (Scientific Basis of Imaging module) and Final FRCR Part A Examination test knowledge through multiple choice and single best answer (SBA) questions. The First FRCR Examination (Anatomy module) tests knowledge by requiring the identification of normal anatomical structures on images. The Final FRCR Part B Examination assesses clinical competence (interpretative, analytical and communication skills)

Formative Assessment

Workplace based assessment will be the cornerstone of assessment for day-to-day practice. There is a range of tools available for this use. These have undergone or are undergoing evaluation in terms of their feasibility, reliability, validity and reproducibility. The generic and radiologically specific workplace based assessment tools are

A. Multisource Feedback

- The multisource feedback (MSF) tool assesses generic skills across the domains of Good Medical Practice. It consists of the collated views from a range of co-workers (previously described as 360° assessment). It will be mapped to a self assessment tool with identical domains
- MSF should usually take place once a year, although the educational supervisor may choose to recommend an additional MSF to investigate a relevant behavioural issue or check progress after an adverse MSF.
- For each assessment, the radiology trainee should nominate 15 raters. A minimum of 12 returns are required.
- Most raters/assessors should be supervising consultants, doctors in training more senior than the trainee under assessment and experienced radiographic, nursing or allied health professional colleagues.

The recommended mix of raters/assessors is

- 2–4 senior doctors
- 2–4 doctors in training
- 2–4 radiographers
- 2–4 nurses/allied health professionals
- 2–4 other team members including clerks, secretaries and auxiliary staff

B. Direct observation of doctor/patient encounter

Three tools can be used to assess radiologist/patient encounters:

- Mini-imaging interpretation exercise (Mini-IPX)
- Radiology- Direct observation of procedural skills (Rad-DOPS)

Radiology trainees are required to undertake a minimum of twelve observed encounters in each year of training although it is anticipated that they may/will undertake many more, as the WpBA are the vehicles by which the trainee will guarantee one-to-one teaching and ensure appropriate curriculum coverage during their clinical attachments.

Mini-imaging interpretation exercise (Mini-IPX)

This is a structured assessment of an observed radiology interpretation/reporting episode:

- trainees should complete a minimum of six mini-IPX in each year of training. These should be spaced out during the year with at least two mini-IPX completed in each four month period.
- a different assessor should be used for each mini-IPX wherever possible, including at least one of consultant level, per four month placement
- assessors **must** be trained in giving feedback and understand the role of assessment.
- mini-IPXs should sample across different clinical radiological problems from the radiology specific content (categories listed in the *Syllabus and Competences* section)
- trainees should agree the timing, problem and assessor.
- assessors may also carry out unscheduled assessments.

Radiology-Direct observation of procedural skills (Rad-DOPS)

This is a structured checklist for assessing the radiology trainee's interaction with the patient when performing a practical procedure:

- trainees must submit a minimum of six Rad-DOPS per annum
- different assessors should be used for each encounter wherever possible
- assessors must be trained both in the procedure and feedback methodology. They could include consultants, more senior doctors in training, advanced practitioner radiographers, qualified nurses or allied health professionals
- Rad-DOPS should sample a wide range of different procedures/skills
- trainees should choose timing, procedure and observer/assessor
- assessors may also carry out unscheduled assessments.

C. Teaching Observation

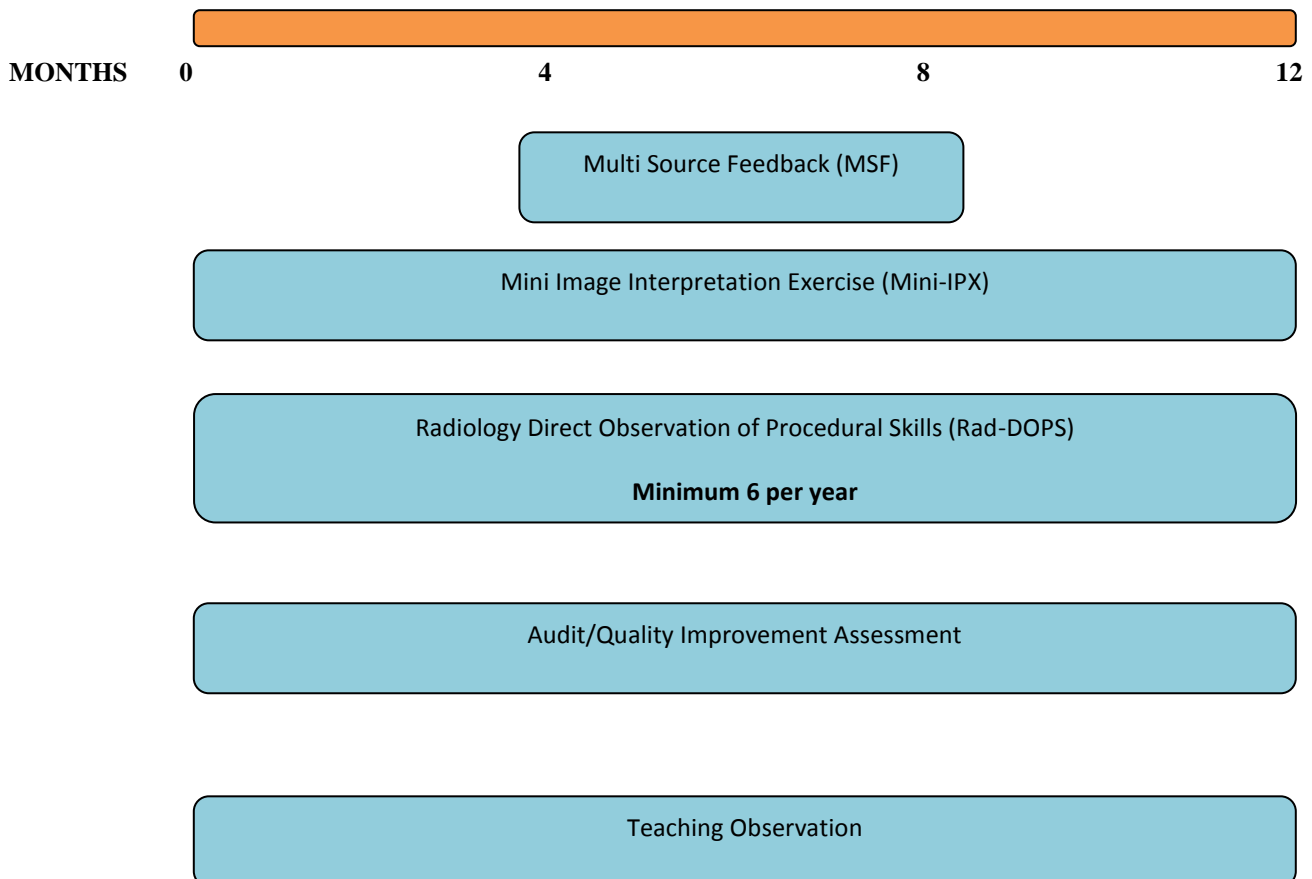
- The Teaching Observation tool evaluates the competence of a trainee to deliver a teaching episode in a wide variety of settings.
- The Teaching Observation form is designed to provide structured, formative feedback to trainees on their competence at teaching.
- The Teaching Observation can be based on any instance of formalised teaching by the trainee, which has been observed by the assessor. The process should be trainee-led (identifying appropriate teaching sessions and assessors).

D. Audit Assessment and Quality Improvement

The Audit Assessment tool is designed to assess a trainee's competence in completing an audit. The assessment can be based on review of audit or quality improvement documentation or on a presentation of the audit at a meeting. If possible, the trainee should be assessed on the same audit by more than one assessor.

All trainees are expected to complete an audit or quality improvement project each year within the training programme. Trainees should show how they have instigated, collated and presented a piece of work, as well as reflected upon any changes in clinical management as a result of work completed.

Figure 1. Assessment during a Year of the Clinical Radiology Training Programme



g. mini-IPX scoring form

mini-Imaging Interpretation Exercise (mini-IPX)

Assessor's Registration Number (e.g GMC, NMC)

--	--	--	--	--	--	--	--

Assessor's Name

Trainee's GMC Number

--	--	--	--	--	--	--	--

--

Date of Assessment (DD/MM/YY)

		/			/		
--	--	---	--	--	---	--	--

Year of specialty training: 1 2 3 4 5

Modality: Plain Film Fluoroscopy Ultrasound CT
 MRI Interventional Radiology Radionuclide Imaging

System: Neuro/ENT Thoracic (CV/Resp) GI/HPB
 Genito-urinary Musculoskeletal Obstetrics/Gynaecology/Breast

Case description:

Setting:

Trainee previous experience of case(s): None Little Average Extensive

Difficulty of case(s): Low Medium High

	<i>Well below expectation for stage of training</i>	<i>Below expectation for stage of training</i>	<i>Borderline for stage of training</i>	<i>Meets expectation for stage of training</i>	<i>Above expectation for stage of training</i>	<i>Well above expectation for stage of training</i>	<i>Unable to comment*</i>
1. Understanding of relevant anatomy	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2. Understanding of clinical context	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3. Infection prevention and control	<input type="checkbox"/> Unsatisfactory <input type="checkbox"/> Satisfactory <input type="checkbox"/> Not applicable						

4. Usage of equipment									
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5. Observation of findings									
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6. Image interpretation									
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
7. Appropriate reference to previous investigations									
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
8. Clarity of report									
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
9. Interaction with patient/staff									
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
10. Judgement/Insight									
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
11. OVERALL CLINICAL JUDGEMENT									
	Rating	Description							
<input type="checkbox"/>	Trainee requires additional support and supervision	<i>Demonstrates little knowledge and lacking ability to evaluate issues resulting in only a minimal contribution to the radiology report and management plan</i>							
<input type="checkbox"/>	Trainee requires direct supervision	<i>Demonstrates some knowledge and limited evaluation of issues resulting in a limited report and management plan</i>							
<input type="checkbox"/>	Trainee requires minimal/indirect supervision	<i>Demonstrates satisfactory knowledge and logical evaluation of issues resulting in an acceptable report and management plan consistent with early higher training</i>							
<input type="checkbox"/>	Trainee requires very little/no senior input and able to practise independently	<i>Demonstrates detailed knowledge and good evaluation of issues resulting in a succinct report and clear management plan</i>							

***Unable to comment** – Please mark this if you have **not observed** the behaviour and feel unable to comment.

Further mandatory questions on the following page

Assessor's comments – state areas of good practice and areas for development *(mandatory field)*

Trainee's comments – comment on your performance and any actions required *(mandatory field)*

Trainee's Signature

Assessor's Signature

h. mini-IPX Guidance for Assessors

The **mini-Imaging Interpretation Exercise** Tool is designed to assess a trainee's skills in interpreting an imaging study and to provide rapid and prompt feedback to a trainee in a particular area of diagnostic imaging.

The trainee should lead the process by identifying the activity to be assessed and appropriate assessors.

Instructions for Assessors:

1. Where the mini-IPX involves patient interaction, please ensure that the patient is aware that the mini-IPX is being carried out.
2. You should directly observe the trainee performing the activity to be assessed in a normal environment.
3. Please assess the trainee on the scale shown. Please note that your rating should reflect the performance of the trainee against that which you would reasonably expect at their stage/year of training and level of experience.
4. Please give an overall rating of the trainee's performance using the options in question 11.
5. Please give feedback to the trainee after the assessment. This should include specific written comments on areas of good practice and constructive feedback on areas for further development.
6. Encourage the trainee to provide comments on their performance and any actions required.

Descriptors of competencies demonstrated during mini-IPX:

Understanding of relevant anatomy

Does the trainee know the relevant anatomical landmarks?

Understanding of clinical context

Does the trainee interpret the images in the full context of supplied clinical information, and seek further information where this is required?

Infection Prevention and Control

The cleansing of hands and, where relevant, equipment before and after every physical patient episode is mandatory.

Usage of Equipment

Does the trainee show an understanding on the radiology equipment with appropriate tool/ probe selection and utilisation? Does he/she optimize equipment parameters for individual examinations?

Observation of findings

Is there recognition of normality and abnormality within the case(s)?

Image interpretation

Are findings interpreted appropriately?

Appropriate reference to previous investigations

Where appropriate

Clarity of report

Does the report have a clear, concise, clinically appropriate and lucid appearance?

Interaction with patients/staff:

Is the trainee able to convey understanding to others?

Judgement/insight

Does the trainee act on and have a sense of urgency when appropriate?
For example, he/she communicates directly with referring clinician to discuss emergent imaging findings

i. Rad-DOPS scoring form

Radiology Direct Observation of Procedural Skills (Rad-DOPS)

Assessor's Registration Number (e.g GMC, NMC)

--	--	--	--	--	--	--	--

Trainee's GMC Number

--	--	--	--	--	--	--	--

Date of Assessment (DD/MM/YY)

		/			/		
--	--	---	--	--	---	--	--

Assessor's Name

--

Year of specialty training: 1 2 3 4 5

Clinical Setting:

Ultrasound Computed Tomography Paediatric Imaging Magnetic Resonance

Imaging

Radionuclide Imaging below)

Interventional Radiology Breast Imaging Fluoroscopy Other (please specify

--

Other setting:

--

Procedure Name:

Number of times this procedure previously performed by trainee: 0 1-4 5-10 >10

Difficulty of procedure: Low Medium High

	<i>Well below expectation for stage of training</i>	<i>Below expectation for stage of training</i>	<i>Borderline for stage of training</i>	<i>Meets expectation for stage of training</i>	<i>Above expectation for stage of training</i>	<i>Well above expectation for stage of training</i>	<i>Unable to comment*</i>
1. Demonstrates understanding of indications, relevant anatomy and technique	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2. Explains procedure/risks to patient, obtains/confirms informed consent where appropriate	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3. Uses appropriate analgesia or safe sedation/drugs	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4. Usage of equipment	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5. Infection prevention and control	<input type="checkbox"/> Unsatisfactory <input type="checkbox"/> Satisfactory <input type="checkbox"/> Not applicable						

6. Technical ability									
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
7. Seeks help if appropriate									
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
8. Minimises use of ionising radiation for procedures involving x-rays									
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
9. Communication with patients/staff									
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
10. Quality of diagnostic images									
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
11. Judgement/Insight									
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
12. Quality of report of procedure									
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
13. OVERALL COMPETENCE									
	Rating	Description							
<input type="checkbox"/>	Trainee requires additional support and supervision	<i>Demonstrates basic radiological procedural skills resulting in incomplete examination findings. Shows limited clinical judgement following encounter</i>							
<input type="checkbox"/>	Trainee requires direct supervision (performed at level expected during Core training)	<i>Demonstrates sound radiological procedural skills resulting in adequate examination findings. Shows basic clinical judgement following encounter</i>							
<input type="checkbox"/>	Trainee requires minima/indirect supervision (performed at the level expected on completion of Core Training)	<i>Demonstrates good radiological procedural skills resulting in sound examination findings. Shows good clinical judgement following encounter</i>							
<input type="checkbox"/>	Trainee requires very little/no senior input and able to practise independently (performed at level expected during Higher Training)	<i>Demonstrates excellent and timely radiological procedural skills resulting in a comprehensive examination. Shows good clinical judgement following encounter</i>							

***Unable to comment** – Please mark this if you have **not observed** the behaviour and feel unable to comment. **Further mandatory questions on the following page**

Assessor's comments – state areas of good practice and areas for development *(mandatory field)*

Trainee's comments – comment on your performance and any actions required *(mandatory field)*

Trainee's Signature

Assessor's Signature

j. Rad-DOPS Guidance for Assessors

The Radiology Directly Observation of Procedural Skills (DOPS) focuses on the skills that trainees require when undertaking a clinical practical procedure. The DOPS is a focused observation or “snapshot” of a trainee undertaking a practical procedure. Not all elements need be assessed on each occasion. You may explore a trainee’s related knowledge where you feel appropriate.

Instructions:

1. Please ensure that the patient is aware that the Rad-DOPS is being carried out.
2. You should directly observe the trainee performing the procedure to be assessed in a normal environment and explore knowledge where appropriate.
3. Please assess the trainee on the scale shown. Please note that your rating should reflect the performance of the trainee against that which you would reasonably expect at their stage of training and level of experience.
4. Please give an overall rating of the trainee’s performance using the options in question 13.
5. Please give feedback to the trainee after the assessment. This should include specific written comments on areas of good practice and constructive feedback on areas for further development.
6. Encourage the trainee to provide written comment on their performance and any actions required.

Descriptors of competencies demonstrated during Rad-DOPS:

Demonstrates understanding of indications, relevant anatomy and technique

Does the trainee know the relevant indications, anatomical landmarks, and techniques relevant to the procedure?

Explains procedure/risks to patient, obtains informed consent where appropriate

Is there a clear explanation of the proposed procedure to the patient, with the patient given an opportunity to ask questions? Where informed consent is sought, is this documented appropriately?

Uses appropriate analgesia or safe sedation

Does the trainee use adequate amounts of appropriate drugs to minimise patient discomfort? Is this titrated where appropriate?

Usage of Equipment Does the trainee show an understanding on the radiology equipment with appropriate tool/ probe selection and utilisation?

Does he/she optimise equipment parameters for individual examinations?

Infection prevention and control

The trainee demonstrates good aseptic technique where appropriate with demonstration of principles of infection prevention and control.

Technical ability

Most pertinent to practical applications such as ultrasound and screening. Is there satisfactory hand/eye co-ordination?

Seeks help if appropriate

Does the trainee recognise his/her limitations and request assistance when appropriate?

Minimises use of ionizing radiation for procedures involving x-rays

Where the procedure involves ionising radiation.

Quality of Diagnostic images obtained

The trainee tailors the number and quality of images to the procedure and patient.

Communication skills with patient/staff

Is the trainee polite, and exhibits a sense of self within a team structure? Is he/she able to convey understanding to others?

Quality of report of procedure

Does the report have a clear, concise, clinically appropriate and lucid appearance, within the context of other available clinicoradiological information?

Judgement/insight

For example, the trainee stops the procedure if unforeseen complications are encountered.

k. Data Coding System

Type	Name	Name	Name
Tree Node	Attitudes	Assessors	Preferences
	Type	personal views	Prior Assessor
	Tree Node	Type	Impression
		Tree Node	
	Tree Node	Authenticity of assessment	
	Tree Node	Influence of number of WPBAs	
	Tree Node	Inhibitory Influences	
		Type	Name
		Tree Node	Trainees
		Tree Node	Trainers
	Tree Node	Prospective vs retrospective	
	Tree Node	Purpose of WPBAs	
		Type	Name
		Tree Node	Formative vs summative
		Tree Node	Pass Fail
	Tree Node	Reliability	
	Tree Node	Subversion	
		Type	Name
		Tree Node	By Trainees
			Type
			Tree Node
			Case Selection
			Tree Node
			Choice of Assessor
			By Trainers
	Tree Node	Tree Node	
		Trainee	
		Motivation	
	Tree Node	Variability of Scoring	
	Tree Node	Who chooses	

		cases and why	
Tree Node	Complex Procedures		
	Type	Name	
	Tree Node	Assessment of Complications and Unexp Events	
	Tree Node	Assessment of teamwork and leadership	
	Tree Node	Assessment of technical steps	
	Tree Node	How to improve assessment	
	Tree Node	Possible for WPBA's	
		Type	Name
		Tree Node	Not Possible
Tree Node	Documentation		
Tree Node	Formalisation		
Tree Node	Good Points of WPBA		
Tree Node	Interviewee details		
Tree Node	Long Term Effects		
	Type	Name	
	Tree Node	Dealing with poor performers	
	Tree Node	Effect on Future Learning	
	Tree Node	Formative usage	
	Tree Node	Relationship with other assessments	
		Type	Name
		Tree Node	End of attachment assessment
	Tree Node	Summative Usage	
	Tree Node	Trust in Results	
Tree Node	Miscellaneous		
Tree Node	Practical Issues		
	Type	Name	
	Tree Node	Assessors Role	
	Tree Node	Choice of Assessor	

	Tree Node	Feasibility	
	Tree Node	Numbers of WPBAs Done	
		Type	Name
		Tree Node	Which types
	Tree Node	Practical Issues	
		Forms	
	Tree Node	Practical Issues	
		Tools	
	Tree Node	Previous	
		Experience of WPBA	
		Type	Name
		Tree Node	Other specialties
	Tree Node	Process and	
		Forms	
	Tree Node	Standard expected	
	Tree Node	Trainees Role	
	Tree Node	Training in WPBA	
Tree Node		Reflections and	
		Feedback	
		Type	Name
	Tree Node	Feedback on	
		Assessor	
	Tree Node	Feedback on	
		Trainees	
	Tree Node	How feedback	
		delivered	
	Tree Node	How results of	
		feedback used	
	Tree Node	Importance of	
		feedback	
	Tree Node	Refection by	
		Trainees on	
		WPBAs	
	Tree Node	Reflection on	
		Assessor of	
		Trainee perf	
	Tree Node	Tickbox vs Free	
		Text	
	Tree Node	Timing of	
		Feedback	
Tree Node		Relationships	
		Type	Name

	Tree Node	Effects on Dep Working
	Tree Node	Effects on Prof Relationships
Tree Node	Standardisation	
Tree Node	Suggested Improvements	
Tree Node	Timing of Assessments	
Tree Node	Trainee confidence and morale	
Tree Node	Weaknesses of WPBA	