

**Centre for Professional Legal Studies
Cardiff Law School**

**Legal Services Commission
'Quality Assurance for Advocates'**

by

Angela Devereux, Jason Tucker, Professor Richard Moorhead and
Professor Ed Cape

Acknowledgments

We are deeply indebted to the assessors and advisers who have ensured that this difficult task has been achieved efficiently, fairly and with some enjoyment - even by the candidates. It would be invidious to single out individuals from the judiciary, but we thank all those judges who assisted us in any way, large or small. We are also grateful to Stephen Wooler and his CPSi team and Chief Crown Prosecutor Chris Woolley for their advice. Special thanks to Northumbria University for opening their Law School to us on a Saturday – the courtroom was wonderful. There are many individuals who have given more time to realising this project than it was fair to expect. Two in particular whose work has been carried on in the wings rather than the spotlight are Diane Davies and Graham Booth, for whom the greatest thanks will be to see the project safely concluded. But we must most of all express our very deepest thanks to all the candidates for their generosity in time and in spirit in engagement in this project.

Contents

1	Introduction	7
1.1	The Research Organisation.....	7
1.2	The Scope of the Pilot	7
1.3	The Pilot Stages	9
1.4	Preliminary Observations	9
2	Methodology.....	10
2.1	History and inheritance of standards	10
2.2	Variety of assessment instruments considered	10
2.3	Assessment instruments compared.....	12
2.4	Assessment instruments and Levels	16
2.5	Assessment instruments and competences	19
2.6	Assignment of advocates to Levels	22
2.7	Dealing with data – preserving confidentiality and objectivity	24
3	Development of the pilot assessment scheme	25
3.1	Assessment development and pre-pilot.....	25
3.2	Standards - cross level performance	26
3.3	Checking assessments in the pilot	27
3.4	Assessment criteria	30
3.5	Finalising data	31
4	Cohort	32
4.1	Background	32
4.2	Cohort make up - general.....	32
4.3	The cohort	33
5	Results from the assessment mechanisms.....	36
5.1	Level 1	36
5.2	Level 2.....	43
5.3	Level 3.....	47
5.4	Level 4.....	50
6	Judicial Evaluation.....	52
6.1	Background to Judicial Involvement	52
6.2	Engagement	53
6.3	Judicial evaluation of all candidates	54
6.4	Comparing Judicial Evaluation with other assessment mechanisms.....	56
6.5	Judicial Evaluation in an operational scheme	57

7	Comparison of different professional groupings.....	58
7.1	Assessment results by professional grouping	58
7.2	CPS External Prosecutor Grading.....	59
8	Which assessments for which level? The possible shape of a scheme.	61
8.1	Cross-checking competence assessment	61
8.2	Place of judicial evaluation	61
8.3	Diets of Assessment.....	62
8.4	Levels and movement	66
8.5	Practical problems and suggested solutions	67
9	Passporting, Exemption and Assessment in an operational scheme.....	70
9.1	Level 1	70
9.2	Level 2.....	72
9.3	Level 3.....	73
9.4	Level 4	73
9.5	CPS External Prosecutor Grading.....	73
10	Setting up an operational scheme.....	74
10.1	Establishing a framework for running a QAA scheme	74
10.2	Who is to assess?	75
10.3	The logistics of assessment	75
10.4	Monitoring and ensuring consistency across assessment organisations....	76
10.5	Feedback.....	77
10.6	Dress	77
11	Costs of an Operational Scheme.....	78
11.1	Candidate Fees – Level 1.....	79
11.2	Candidate Fees – Level 2.....	79
11.3	Candidate Fees – Level 3.....	79
11.4	Candidate Fees – Level 4.....	80
11.5	Costs of any reaccreditation	80
11.6	Costs of Regulation	81
12	Summary and conclusions	82
12.1	Numbers participating.....	82
12.2	Choice of assessment instruments.....	83
12.3	Assessment mechanisms by levels	84
12.4	Assignment of advocates to Levels	85
12.5	Testing of simulated assessment mechanisms	86

12.6	Judicial Evaluation	86
12.7	Results from the assessment mechanisms	87
12.8	Level 1	87
12.9	Level 2	88
12.10	Level 3	88
12.11	Level 4	89
12.12	Judicial Evaluation	89
12.13	Results for judicial evaluation of all candidates	90
12.14	Judicial Evaluation in an operational scheme	91
12.15	Which assessments for which level? Recommendations for the implemented scheme	91
12.16	Passporting and associated issues	94
12.17	Equalities and diversity	98
12.18	Setting up an operational scheme	98
12.19	Feedback to candidates, appeals, monitoring and ensuring consistency across assessment organisations.....	99
12.20	Costs of an Operational Scheme.....	99

List of Tables

Table 1: Lowest and highest levels at which candidates felt they should be assessed	23
Table 2: Assessed.....	33
Table 3: Participation by professional grouping.....	34
Table 4: Attrition by level and professional grouping (percentage either withdrawing or unable to sit assessments).....	34
Table 5: Candidates' Degree.....	35
Table 6: Cross examination level by professional group	38
Table 7: HRA by cross examination result (Level 1)	39
Table 8: Level 1 mechanisms compared.....	41
Table 9: Summary of Level 1 failures.....	42
Table 10: Cross examination compared with examination in chief.....	43
Table 11: Cross examination Level 2 compared with Multiple Choice Test.....	44
Table 12: Written advocacy compared with examination in chief	45
Table 13: Scores on the MCT crossover questions by Level of Assessment.....	46
Table 14: Summary of candidates with at least one fail (Level 2)	47
Table 15: Portfolio and Written Advocacy compared (Level 3).....	49
Table 16: Performance across Level 3 mechanisms.....	50
Table 17: Level 4 Results.....	51
Table 18: Mean scores Judicial Evaluation (assessed and evaluated only).....	55
Table 19: Percentage failure rate of Live Day assessments by professional grouping (Level 1)	58
Table 20: Percentage failure rate of Live Day assessments by professional grouping (Level 2)	58
Table 21 Assessment Level numbers by CPS (EP) grading	59
Table 22: CPS Cross examination Level achieved compared with CPS (EP) grading	60

1 Introduction

The impetus for the Quality Assurance for Advocates (QAA) pilot was provided by Lord Carter's Review of Legal Aid Procurement (July 2006), which noted that, whilst there were quality assurance mechanisms in place for legal advice, assistance and litigation, there was little quality assurance of advocacy other than reactive, complaints based mechanisms and traditional professional training and entry regimes. Part of recommendation 5.3 of Lord Carter's report advocated:

"A proportionate system of quality monitoring based on the principles of peer review and a rounded appraisal system should be developed for all advocates working in the criminal, civil and family courts."

Recommendation 5.3 also noted that "the new quality monitoring system should be developed in the first instance for publicly funded criminal advocates". The focus of the QAA pilot has, therefore, been criminal advocacy.

1.1 The Research Organisation

The Research Organisation chosen to conduct the QAA pilot was Cardiff University. The Research Team consisted of staff from Cardiff Law School: Angela Devereux (Principal Investigator), Professor Richard Moorhead and Jason Tucker; and Professor Ed Cape (University of the West of England).

The Research Team was assisted by a group of advisers who were consulted regarding the assessment instruments which should be used during the pilot. The advisers included a retired circuit judge, a Chief Crown Prosecutor, a clerk to the justices and a pool of criminal practitioners.

The assessments were conducted by a team of assessors consisting of three academic staff from Cardiff Law School's Centre for Professional Legal Studies and seven current practitioners. The academic assessors were all former barristers or solicitors and experienced advocates and advocacy teachers. They included a former CPS grade 3 prosecutor and current recorder. The practitioner assessors were all experienced advocates, including two Queen's Counsel, and the majority of them had considerable prior assessment experience, primarily undertaking assessments for the Criminal Litigation Accreditation Scheme (Court and Police Station for Duty Solicitors).

1.2 The Scope of the Pilot

The original specification for the QAA pilot required us to:

- research, analyse and report on assessment options that can be used to effectively assess advocates against the defined competence framework;
- make recommendations to the QAA Project Team as to the most effective assessment route that best covers the 4 levels of advocacy to be tested in

the pilot, to include the range of assessment methodologies that may be employed to cover sections A – E of the competence framework;

- consider alternative options that may be necessary in order to provide flexibility according to different types of practice and to ensure equality of opportunity for all advocates;
- define the range of evidence that may be employed to demonstrate competence at each level;
- map all appropriate existing learning and accreditation programmes applicable to crime barristers and solicitor advocates and make informed recommendations on the scope for complete passporting or partial accreditation within the QAA Scheme.
- It was envisaged that we would conduct a maximum of 250 assessments in the pilot to test the validity, reliability and cost effectiveness of the assessment process.

We were provided with a competence framework (which appears as Annex A), and also with the advocate levels. These were not developed by the researchers but developed by work stream groups reporting to the Reference Group (see below at 2.1), who signed them off for testing. The Levels are set out in the following grid:

Guidance on QAA Pilot Levels

	Description
Level 1	Magistrates Court, Appeals to the Crown Court and Committals
Level 2	More straightforward Crown Court – e.g. Jury trials including lesser offences of theft, dishonesty, deception and handling, assault (ABH and section 20) burglary (not aggravated), lesser more straightforward drug offences and lesser offences involving violence or damage, plus straightforward robberies and non fatal road traffic offences. Also, sexual offences and less serious offences against children
Level 3	More complex Crown Court and above – More serious cases of dishonesty and fraud. Drug offence such as possession with intent to supply drugs, blackmail, aggravated burglary, violent disorder, arson, complex robberies, serious assaults, driving offences involving death, child abuse and sexual offences under the Sexual Offences Act 2003, plus serious sexual offences.
Level 4 Inclusive of QC's	The most complex Crown and High Court cases. Very serious, sensitive and complex cases, including serious sexual offences, substantial child abuse, very serious and multi handed murder trials, cases involving issues of national security, serious organised crime, terrorism and complex and high value frauds.

The original scope of the pilot was to consider publicly funded defence advocates. However, by agreement with the Crown Prosecution Service (CPS), the scope of the pilot was extended to include up to 30 advocates from the CPS.

1.3 The Pilot Stages

The pilot was divided into three Stages:

Stage 1 - Assessment Research and Design

The principal objective for Stage 1 was to research and report on options for assessment of the key skills identified in the QAA competence framework.

Stage 1 was commenced in October 2008. The progressive nature of the scheme meant that aspects of research continued until May 2009.

Stage 2 - Assessment Testing

The principal objective for Stage 2 was to conduct pilot assessments across the four proposed levels of advocacy and generate, collect, record, collate, store and review data to support the evaluation of the QAA pilot.

Stage 2 was conducted between March and September 2009.

Stage 3 - Evaluation

The principal objective for Stage 3 was to analyse the quantitative and qualitative data generated during the pilot and to report to the Commission in respect of the key matters identified in the project specification. It is the Commission's intention to use the findings from the QAA pilot to inform a subsequent consultation regarding quality assurance for publicly funded advocates.

Stage 3 was conducted in October and November 2009.

1.4 Preliminary Observations

The original specification for the QAA pilot envisaged that drawing on data from approximately 280 assessments (including the additional CPS candidates) we would map existing accreditation programmes and make recommendations on the scope for complete passporting or partial exemption within any QAA Scheme. It was also necessary for the pilot stage of the contract to contain the split of professional groupings, with the relevant types of accreditation/experience. Our aims in this regard are set out in Annex E.

However, the number of advocates who actually participated in the pilot was 101, of whom 98 completed assessments in time for evaluation. Whilst these were spread across the 4 levels and came from different professional groups, the numbers participating has significantly reduced our ability to robustly test the feasibility of potential passporting arrangements.

2 Methodology

2.1 History and inheritance of standards

Some 2 years before the contract for QAA was awarded, a reference group consisting of members of all professions, academics and stakeholders, had undertaken considerable work creating the competences, levels and criteria which they hoped would be used in a scheme for quality assurance for advocates. When their paper - a joint Ministry of Justice (MoJ) and Legal Services Commission (LSC) document - went out for consultation in 2007 it became clear that there were aspects of those structures which required further testing and research.

Thus when the research team started work on the project it inherited some standards - Competences (Annex A) and Levels (Annex B and section 1.2) - to be tested. An incarnation of the working group remained in existence and has received reports on the project, with a smaller dedicated committee (the Assessment Work Stream Group) given the task of both advising the research team as appropriate and reporting back from those meetings to the main group. This liaison has acted as a useful touchstone at key points in the design part of the project and means that there remained a direct line of communication to those who had undertaken the initial work.

Additionally throughout the project there has been frequent contact in progress meetings and otherwise with the QAA team from the LSC, as well as meetings with MoJ representatives, with senior members of the judiciary, the CPS and the CPS Inspectorate.

The project has been subject to Cardiff Law School's research ethics procedures.

2.2 Variety of assessment instruments considered

The research team met early in the project to review the various types of assessment instrument in common use and to consider their utility and relevance for assessing the competences of advocates. They had regard to existing modes of assessment used by both sides of the profession, and by the CPS, in drawing up the framework of assessments for consideration. The methods considered were assessment of:

- the performance of an advocate in a real trial (by a professional assessor and/or the trial judge);
- an advocate's conduct of a real trial by an advocate's own description of that involvement in a portfolio case by an assessor or review from a file;
- an advocate's written advocacy, anonymised from a real case;
- the performance of an advocate in a simulated hearing by an assessor;
- a written examination with problem questions or multiple choice questions;
- and,

- an application made by an advocate supported by references.

From the above list we identified those instruments which could provide **reliable, verifiable and consistent evidence** which would be of research value and would also be able to form **part of an operational scheme**. As a result, the following methods of assessing were rejected:

2.2.1 The performance of an advocate in a real trial by an assessor

Whilst able to be an objective assessment made according to agreed standards this would not be capable of comparison or verification for appeal. Earlier research on public defenders conducted by members of the research team had shown significant economic and practical problems in using trained assessors to assess advocates in real trials.¹ The trials may not be effective, will vary in substance and involvement by the advocate, could be subject to relisting without notice or may, late in the day, be handed to an advocate other than the one whom it is sought to assess. Assessors would face significant difficulties in scheduling assessment visits and would face high levels of wastage when trials did not materialise as planned. CPS inspectorate assessors did assess real hearings for their own 2009 advocacy report, but many of the hearings which the CPSi were content to observe were not trials (thus could not enable testing of many of the competences forming the QAA framework) and even then they achieved only 25% time efficiency. They did not face the additional (and considerable) restriction of needing to report on the performances of specific advocates.

2.2.2 Review from a file

File review was rejected because of the difficulty of assessing advocacy by the kinds of documentation which would appear on a file. Barristers faced the further difficulty of not generally having files for cases once instructions were completed. In solicitors' firms, files are often the work of a number of people, thus preventing individual assessment. This method too failed on both counts.

2.2.3 An application made by an advocate supported by references

An application made by an advocate and supported by references could not give verifiable evidence for assessment purposes. Unlike a case report in a portfolio, an application provides only general information about the nature of an advocate's work. Because of this the danger of putting a favourable 'spin' on a practice is increased, due to the inability of any assessor to look for evidence to verify or refute the assertions made. It was also likely to have posed operational problems, due to difficulties in advocates' obtaining the necessary supporting references.

¹ L. Bridges, E. Cape, P. Fenn, A. Mitchell, R. Moorhead and A. Sherr, Evaluation of the Public Defender Service in England and Wales (The Stationery Office, London 2007).

2.3 Assessment instruments compared

We sought to test a range of different assessment instruments with a view to making recommendations about possible assessment regimes from within that range. A full armoury of assessment instruments would therefore give the best information to be able to make recommendations about which one, or which combination, would give the best assurance about an advocate's competence.

It was hoped that trialling a variety of assessment methods would thus allow the team to pare down the assessments needed to show the relevant competences. Where more than one assessment for a single individual arrived at results confirming one another, then it would be considered whether one could be proxy for the other, or if both (all) were necessary, by looking to any parts of the competence framework (which is set out in detail at 2.5 below) which one tested but the others could not.

Assessing an advocate's performance in a simulated situation by reference to a single or several case studies could allow a number of competences to be tested if the advocate were required to perform the different tasks upon which good advocacy is founded. Additionally some assessment methods commended themselves at particular levels because they would tie in with existing qualification assessment regimes. They would thus be more readily accepted and could, if the professions so desired, take the place of those regimes – thus avoiding a multiplicity of qualifications.

The assessment methods tested can be categorised as follows:

- simulated (advocacy and other) assessment methods;
- assessment methods deriving from real cases;
- assessment of advocacy in real cases.

2.3.1 Simulated (advocacy and other) assessment methods

Methods used in a controlled situation with cases or questions common to sets of advocates and written by the assessment team:

- Interview/Conference (I/C);
- Submission (i.e. non-witness based advocacy) (Sub);
- Cross examination (XX);
- Examination in chief (XC);
- Multiple Choice Test (MCT);

Submission, cross examination and examination in chief are referred to generally later as Live Advocacy (LA).

For the Live Advocacy assessments, candidates were provided with case documentation and asked to prepare for the particular activities which they then performed in front of one or more assessors. Performances were recorded.

The Live Advocacy exercises look most directly at the core skills of the criminal advocate. Simulations are a proxy for live assessment of real cases (such as observation by assessors in real court situations). They have some limitations. Notably, candidates may improve their performance for an assessment (performing to a standard which is higher than their normal standard). Conversely, assessment-based nerves, or the unreality of simulations, may lead them to underperform. A key advantage of simulation over real life assessment is that the simulations ensure that all candidates work with the same materials and are tested on the same facts and issues. Thus an assessment of one candidate is broadly comparable with the assessment of another candidate.

The simulation also allows a wider range of issues and tests to be built in than would ordinarily be found in a real case. The simulated advocacy exercise can test the performance skills of the advocate and their ability to identify and use evidential and legal points. With suitably designed assessments, it provides the opportunity to feed the candidate late information to test the speed and accuracy of their response. It also tests their ability to adapt 'on their feet' to a response not in line with expectations, or a witness's particular characteristics. Thus, in important senses, the 'unreality' of simulation can be an advantage: it can be designed to include particular characteristics essential for the testing of competence at the relevant level. It could take a number of real cases to provide an opportunity to test the full range of competences that are testable in a simulation.

In common with the Live Advocacy assessments, limited papers for the interview were provided in advance to the candidates. These were supplemented on the assessment day by additional paperwork and by instructions from the client (played by an actor). The interview assessment gives an opportunity to assess the advocate's interaction with the client (competences in section C of the framework), as well as their ability to respond to new material. It also gives the opportunity, when used in conjunction with the submission (which derived from it), to assess the advocate's ability to make only relevant submissions (competence A.1.3) and to develop arguments in a logical order (competence A.2.3). Its specific utility in the context of the competence framework is in assessing client interaction. Thus, unless linked with a submission, it is a resource intensive way of providing information on a limited, though important, part of the framework.

The exercises were developed by the research team in consultation with the assessors. The risk of the cases used in simulated oral advocacy becoming known over time was carefully guarded against. Each candidate entered into an agreement with the research team as to the way in which they would treat the papers to ensure the assessments of other candidates were not compromised (see Annex C). Papers were sent in prominently marked and sealed envelopes within other envelopes containing suitable letters of warning about the confidentiality of the contents. Unused papers were returned to us or shredded. All papers and

notes were collected from the candidates at the end of their Live Assessment Day. We have no reason to believe from behaviour on assessment days, or from the levels of performance achieved, that candidates shared information about the simulations.

Multiple Choice Tests were included as a test of the range of each candidate's knowledge, appropriate to the level at which they were being assessed. A strength of Multiple Choice Tests is that they can provide a wide assessment of the coverage of a candidate's knowledge, and can include (as did the tests used here) particular issues which a candidate 'must-know' if they are to function successfully as a criminal advocate.

A disadvantage of Multiple Choice Tests is that they provide only a limited indication of how well the candidate can apply and manipulate that knowledge. Another disadvantage is that they may falsely indicate weakness because an advocate might expect in practice to be able to look up the answers to many of the kinds of questions which could be posed in an Multiple Choice Test, and will not therefore have committed them to memory.

No texts were allowed for this test. To permit this would have changed and sometimes devalued the evidence gained from the answers. Some gaps in knowledge or understanding can seriously compromise the judicial process, requiring us to test candidates' ability to instantly recall such knowledge. Our recognition that in reality an advocate might, without prejudice to the client, have been able to look up some answers resulted in our refining the data eventually derived from the Multiple Choice Test to enable us to focus on the questions we would expect them to be able to answer 'on their feet' (the 'must-know' questions are identified in the appendices). The MCT was taken on the same day as the Live Advocacy in a controlled situation.

2.3.2 Assessment methods deriving from real cases

Methods used to test accounts or parts of real cases written up or drafted by the advocate and assessed by common criteria devised by the assessment team:

- Portfolio (Pf);
- Written Advocacy (WA).

Because it was not, in our view, possible cost-effectively to assess real performance in courts other than through judicial evaluation (see below) other methods were used which relied on practitioner reports to test accounts or parts of real cases written up or drafted by the advocate. Such methods provide some indication of the level of experience and competence of an advocate, but are subject to greater presentational manipulation and they may also test candidates' abilities to understand what the assessors want to see, as much as they test their actual competence.

The use of a portfolio of some of the advocate's cases as an assessment instrument has the advantage of providing a "slice" of the candidate's real practice.

It requires the candidate to produce a document which sets out certain specified aspects of cases they have dealt with, in order to satisfy stated criteria. The candidates are told the format for the document and the criteria against which it will be marked. They are asked to produce a specified number of cases of a certain type. To the extent that it is in the nature of a portfolio that candidates produce their “better” cases it gives some sense of the breadth and depth of their practice.

So, for example, a portfolio at Level 1 requiring two trials to be submitted would include a brief initial description of the nature of each case and the nature of the advocate’s involvement. It would then be necessary to describe in detail: the parts of the process required to enable judgment under the criteria e.g. setting out the nature of the evidence; the client’s instructions and any change in them; the advice given and whether it was taken; any particular difficulties which the client, the nature of the case or the witnesses posed and the approach taken to the gathering or undermining of any evidence.

The document seeks to enable an assessor to judge the extent to which the relevant competences are displayed and whether at an appropriate level. The document is expressed in such a way as to preserve client confidentiality at all times, together with candidate anonymity until the marking process is completed. It has the advantage of being a format which is well understood, tested and accepted amongst many solicitors having been the constant assessment instrument for showing competence as a police station and court duty solicitor. It is also a type of instrument for which there are already sets of assessment criteria enable insight into an advocate’s practice.

It has the following disadvantages:

- For those whose practice is good but who do not put sufficient effort into its preparation it will give a result which does not reflect their level of practice. Good guidance reduces this risk; as does the recognition by candidates’ senior colleagues or employers of the need to devote time to its preparation.
- For those whose practice is less strong, it gives the opportunity to massage the account of that practice to give the best possible “spin” on the candidate’s performance. Its use in conjunction with other assessment instruments guards against this.
- For those whose practice, due to reasons outside their control, is not at as high a level as their skills, it may not allow them to show the true extent of their abilities. Its use in conjunction with other assessment instruments can reduce the impact of this. It can be further mitigated by a scheme which allows flexibility between Levels.

In the written advocacy test candidates were asked to provide a suitably anonymised piece of written advocacy. This provided an opportunity to assess the extent to which an advocate can research, construct and present a legal argument. Such an argument could derive from a trial in the portfolio, or from another case and, because written advocacy is increasingly a part of the armoury of an

advocate, it is appropriate to test competence in this form. It is an assessment of the candidate's real competence rather than their potential, but is nonetheless (unlike their actual performance in a trial) capable of moderation and detailed comparative scrutiny. It is the only instrument which allows the assessor specifically to mark the candidate's written fluency and use of language.

It would be possible to assess written advocacy by virtue of a simulated exercise. One might argue that to do so would have similar benefits as for simulated advocacy, namely consistency and ability to test a prescribed range of knowledge. However, it would be more burdensome on the advocate, who would be required to assimilate new facts and draft a new argument instead of submitting one already in existence. The creation of the additional case studies would also make the accreditation process more expensive.

2.3.3 Assessment of advocacy in real cases

The mechanism for assessing practitioner performance in real cases in 'real time' was Judicial Evaluation (JE).

Whilst subject to the criticism that it is not verifiable (in the sense that it would be difficult for a judicial evaluation to be checked and appealed in the same way as other assessments), judges are exposed to advocacy every day that they sit and so have a ready basis for forming judgments on the quality of the advocate before them. This pilot provided an opportunity to trial the practicality of judicial evaluation and, with enough participation from judges, provide an indication of its reliability through enabling us to compare the ways in which different judges scored the same candidates. Furthermore, it was hoped we would be in a position to compare judicial evaluations with results from other assessment mechanisms to assist in testing the reliability of those mechanisms and helping to choose which would be recommended for eventual use in any quality assurance scheme.

2.4 Assessment instruments and Levels

The Levels dividing the advocate candidates, and the standards to which we had to assess them, were provided by the Levels Work Stream Group. We were asked to divide advocates according to the cases they would deal with. Those Levels have been set out in brief already at section 1.2. The full explanatory document prepared by the Levels Group appears as Annex B, but we here rehearse in more detail the grid of offences from that document by which the advocate's work has been defined for the purposes of QAA.

Table of proposed offences

Level 1	Level 2	Level 3	Level 4
Magistrates Court Work	C Lesser offences 'involving violence or damage, and less serious drug offences'	B Offences using serious violence or damage, and serious drug offences	A Homicide and related grave offences
Appeals from Magistrates to Crown Court	D Sexual offences and less serious offences against children	D Sexual offences and offences against children	J Serious sexual offences
	H Miscellaneous other offences	I Offences against public justice	
	E Burglary		
	F Other offences of dishonesty	G Other offences of dishonesty (£>30,000)	K Other offences of dishonesty (definition required)
Committal for sentence	Cases falling in Magistrates jurisdiction but client elect trial by jury: original advocate instructed should continue to represent the client	J Serious sexual offences, K Other offences of dishonesty	Terrorism offences

In deciding which instruments to use at the respective levels the research team took into account:

- the range of competences to be assessed;
- their robustness as tests of competence;
- the relative utility of the respective instruments;
- value of evidence based on numbers;

- practitioner familiarity with the various instruments;
- the feasibility of designing and implementing appropriate instruments;
- the cost of assessment mechanisms to practitioners and commissioners;
- acceptability in an operational scheme.

The assessment regime for the pilot was as follows:

Table showing assessments used for the pilot

Level	Judicial Evaluation	Portfolio	Written Advocacy	Cross examination	MCT	Inter-view	Submission	Examination in Chief
1	—	✓	✓	✓	✓	✓	✓	—
2	✓	✓	✓	✓	✓	—	—	✓
3	✓	✓	✓	✓	—	—	—	—
4	✓	✓	✓	—	—	—	—	—

Because we predicted that Levels 1 and 2 would give rise to the greatest number of candidates, we tested the widest range of mechanisms at this level.

The team decided after consultation that it would be inappropriate to seek to use Judicial Evaluation at Level 1 as it would have been necessary to involve lay magistrates in the assessment regime. This method was however to be used at all other levels.

Cross examination is the skill which allows of testing of more of the stated competences than any other instrument (see 2.5). It was tested at levels 1 to 3, with the team utilising the Level 3 work to test the assessor's confidence in using a similar approach at Level 4.

Evidence of candidates' own practices was also sought at all Levels in the form of portfolio cases and examples of written advocacy.

Having candidates present for an assessment day also afforded the opportunity to check breadth of knowledge in law, evidence and procedure by setting them an MCT. The presence of a jury at Level 2 increases the importance of an advocate's knowledge of evidence and procedure. The best test of this knowledge across a wide range of scenarios was felt to be the MCT. Given the limitations of MCTs in assessing higher levels of skill, this was targeted at candidates for Levels 1 and 2.

As note above, the added value of an interview assessment is in assessing client-related skills. Because interviewing is part of the assessment regime for Duty Solicitors it was felt appropriate to target the assessment of this skill at Level 1 candidates. It was also felt to be inappropriate to expect higher level candidates to

undergo this assessment, given the extra demands of the live assessments beyond Level 1. In an operational scheme an advocate at or beyond Level 2 would have ordinarily gone through the earlier assessment of interview skills at Level 1.

2.5 Assessment instruments and competences

The competences for an advocate had been drawn up by the Competences Work Stream Group prior to the awarding of the contract. They were those which, save for those at E (which only applies to the leading of cases by senior counsel), are common to and necessary for good advocacy at all levels. They are capable of fulfilment and assessment at all levels. It was therefore important in designing the assessment instruments that they not only reflected these competences, but also that the research team were aware of the instruments capable of assessing each one. This was to be of significance in identifying any untested competences but also in ensuring that any reduced diet of assessments still covered them.

We therefore analysed the competences in order to identify which of them could be assessed in which assessment instrument(s). That analysis follows.

Where the instrument is in bold below alongside the competence, it indicates that this competence will most readily be capable of evaluation within this instrument; where the instrument is not in bold a competence may arise for judging under an instrument, but there is no certainty of that.

A: ANALYSIS

A.1. Accurately identifies key legal and factual issues

- | | | |
|---|------------------------------------------------|-----------------------------|
| 1 | Presents and questions only material witnesses | Pf |
| 2 | Asks only relevant questions | LA and JE |
| 3 | Makes only relevant submissions | Pf, WA; JE; Sub; MCT |

A.2. Adopts appropriate structure and sequence

- | | | |
|---|----------------------------------------------------------------------------------|----------------------|
| 1 | Has a clear strategy for the case | Pf; LA and JE |
| 2 | Case strategy is supported by questions asked and evidence called (see D2 below) | Pf; LA and JE |
| 3 | Develops arguments in a logical order | WA and Sub |

A.3. Responds appropriately to new evidence

- | | | |
|---|-------------------------------------------------|--------------------|
| 1 | Makes appropriate objections and/or submissions | Pf; JE; MCT |
| 2 | Asks appropriate questions | LA and JE |
| 3 | Takes appropriate advantage of new material | Pf and LA |

B: ORGANISATION**B.1. Is thoroughly prepared**

- | | | |
|---|-----------------------------------------------------------------|------------------|
| 1 | Understands opponent's case and assimilates opponent's evidence | LA and WA |
| 2 | Locates materials and evidence quickly | LA and JE |
| 3 | Assists the court where consistent with duty to the client | Pf and JE |

B.2. Observes procedures

- | | | |
|---|--------------------------------------------------------------------|---------------------------------------------------------|
| 1 | Complies with appropriate Procedural Rules and judicial directions | WA; JE; MCT |
| 2 | Provides appropriate disclosure of evidence | Pf (if arises)
more applicable to prosecutors |
| 3 | Obtains instructions when appropriate | Pf |

B.3. Meets deadlines

*This whole section incapable of assessment **routinely** except by portfolio and possibly JE.*

- | | | |
|---|--------------------------------------------------------|-------------|
| 1 | Keeps the court informed of any timing problems/delays | Pf |
| 2 | Complies with judicially imposed timetables | Pf |
| 3 | Is punctual | None |

C: INTERACTION**C.1. Assists client in autonomous decision making**

- | | | |
|---|------------------------------------------------|--------------------------------------------------|
| 1 | Gives lay and professional client clear advice | I/C and Pf |
| 2 | Keeps lay and professional client up-to-date | Pf
(if a self assessment box provided) |
| 3 | Ensures lay client understands the process | I/C and Pf |

C.2. Establishes professional relationships in court

- | | | |
|---|--------------------------------------------------------------|-----------------------------|
| 1 | Observes professional etiquette in relation to third parties | P/f and JE |
| 2 | Is courteous at all times | LA and JE |
| 3 | Keeps all parties informed | Pf (see note at end) |

C.3. Respects witnesses

- | | | |
|---|-----------------------------------------------------------|-----------------------------------------------------------------|
| 1 | Observes restrictions and judicial rulings on questioning | Pf and JE |
| 2 | Deals appropriately with vulnerable witnesses | Pf
(only if criteria made it a technical requirement) |
| 3 | Deals effectively with uncooperative witnesses | LA (if so designed) |

D: PRESENTATION**D.1. Presents clear and succinct written and oral submissions**

- | | | |
|---|----------------------------------------------------------|--------|
| 1 | Drafts clear skeleton arguments | WA |
| 2 | Speaks clearly and audibly | LA |
| 3 | Maintains pace <i>throughout the course of the trial</i> | JE and |
| | without italicised part, assessable through | LA |

D.2. Conducts focused questioning

- | | | |
|---|------------------------------------------------------------|---------------|
| 1 | Questions to witnesses are clear and understandable | LA |
| 2 | Questioning strategy is clear and relevant to issues | LA; Pf and JE |
| 3 | Avoids introducing irrelevant matters in cross examination | LA and JE |

D.3. Observes professional duties

- | | | |
|---|-------------------------------------------------------------------------------------------|----------------|
| 1 | Observes duty to the court and duty to act with independence | Pf and JE |
| 2 | Advises the court of adverse authorities and, where they arise, procedural irregularities | PF; JE and MCT |
| 3 | Assists the court with the proper administration of justice | Pf; JE and MCT |

E: LEADING CASES

Range: Leading in a complex case (Levels 3 and 4 only)

E.1. Effectively leads an advocacy team

- | | | |
|---|------------------------------------------------------------------------------------|-----------|
| 1 | Takes responsibility for effective case management | Pf and JE |
| 2 | Ensures team members are allocated tasks consistent with their level of competence | Pf |
| 3 | Effectively demonstrates ultimate responsibility for the case | Pf and JE |

Competences **B.3.1. Keeps the court informed of any timing problems/delays;**
B.3.2 Complies with judicially imposed timetables and **B.3 3 Is punctual;**
C1.2 Keeps lay and professional client up-to-date and **C 2.3 Keeps all parties informed** all deal with timing or the passing of information about timings. Whilst some of these (B.3.1-3.3) may be assessable in part through judicial evaluation, without significant (and currently unfeasible) monitoring of practice, these cannot otherwise be assessed except by means of self-assessment. This is unlikely to be rigorous or helpful.

We recommend that competences B3 (1, 2 and 3); C1.2 and C 2.3 be removed from the Competence Framework

Similarly, **C.2.1 Observes professional etiquette in relation to third parties** is not capable of definite assessment under any QAA scheme that the research team can envisage except by self-certification. Being neither rigorous nor helpful:

We recommend that C.2.1 be removed from the framework

Additionally there are competences which we deem to be sure of assessment by **only one** instrument:

A 1.1 Presents and questions only material witnesses	Pf
A.3.3 Takes appropriate advantage of new material	Pf and LA
B.2. 2 Provides appropriate disclosure of evidence	Pf
(if it arises, more applicable to prosecutors)	
B.2. 3 Obtains instructions when appropriate	Pf
C.3. 2 Deals appropriately with vulnerable witnesses	Pf (only if criteria made it a technical requirement)
C.3. 3 Deals effectively with uncooperative witnesses	LA (if so designed)
D.1.1 Drafts clear skeleton arguments	WA
D.1.2 Speaks clearly and audibly	LA
D.1.3 Maintains pace <i>throughout the course of the trial</i>	JE (but without italicised part, assessed also in LA)
D.2.1 Questions to witnesses are clear and understandable	LA
D.3.3 Assists the court with the proper administration of justice	MCT
E.1.2 Ensures team members are allocated tasks consistent with their level of competence	Pf

We will revert to this list when making recommendations about the regime of assessments desirable in an operational QAA scheme.

2.6 Assignment of advocates to Levels

When volunteering for the pilot, candidates were given the summary of Levels which appears at section 1.2 and the document prepared by the Levels Group (Annex B). Amongst the data sought from the candidates was data about the levels of cases which they actually conducted. Some candidates entered only one level. These candidates, if available for assessment, were assessed at that level.

Many candidates entered two levels. We then decided the level at which they ought to be assessed based on any other information provided by the candidate such as years since call or admission, or other experience gained. The likelihood was that those indicating both Levels 1 and 2 would be invited for a Level 2 assessment. Candidates sometimes contacted us and provided information which led to reassessment of the level at which they should be assessed. For example, some who had appeared in the Crown Court but only in respect of appeals from a magistrates' court were reassigned to Level 1. The same allocation exercise took place at Levels 2 and 3 when both Levels were indicated by the candidate.

Those who ticked both Levels 3 and 4 were (if they were available for a Live Assessment Day) assessed at Level 3. This was because the written part of their assessment (a single case portfolio and written advocacy) was the same for both Levels. The only differentiator of Level for this document was the level of case chosen by the advocate for description in the portfolio and the written advocacy. As

there was to be no “live” assessment at Level 4, candidates indicating 3 and 4 would thereby be assessed over a wider range of instruments than would otherwise be the case. In any event none of the candidates who had so indicated was ultimately available for live assessment.

Where candidates indicated three levels we registered them as the middle one – but were able to modify it following a request and explanation. Where a candidate put in all four levels we formed our own view. In the event, none of the candidates doing that made themselves available for assessment.

The fact that candidates found it difficult to categorise with any certainty either their own level or that of the cases with which they mostly dealt is an important indicator of potential problems in using the Levels as drafted.

Candidates were asked to indicate the Level(s) at which they felt they should be assessed. Table 1 shows the results of that. Because candidates indicated a range of levels the highest and lowest levels indicated by them are shown in the table. About half of those indicating they could be assessed at Level 1 or Level 2 indicated they could also be assessed at a higher level.

Table 1: Lowest and highest levels at which candidates felt they should be assessed

			Highest Level				
			1	2	3	4	Total
Lowest Level	1	Number	29	18	2	0	49
		%	59.2%	36.7%	4.1%	.0%	100.0%
	2	Number	0	17	14	0	31
		%	.0%	54.8%	45.2%	.0%	100.0%
	3	Number	0	0	10	3	13
		%	.0%	.0%	76.9%	23.1%	100.0%
	4	Number	0	0	0	5	5
		%	.0%	.0%	.0%	100.0%	100.0%
	Total	Number	29	35	26	8	98
		%	29.6%	35.7%	26.5%	8.2%	100.0%

We can also compare their assessments with our assessments.

- 60 indicated one level and were assessed at that level;
- 17 indicated a range and were assessed at the higher end of the range (this included people assessed at Levels 2, 3 or 4);
- 17 indicated two levels and were assessed at the lower level;
- 1 candidate indicated they were assessable at a Level 2 but was in fact assessed at Level 1;
- 1 indicated they were a 3 or a 4 and were assessed at Level 2; and,
- indicated a range of three levels and were assessed at the mid level (Level 2 on both occasions).

For a scheme based on these Levels to be efficient when operational, it would be necessary for the advocate, their clerk (where relevant), client and or the paying body to be clear about the Level of the case for which public funds were being sought.

2.7 Dealing with data – preserving confidentiality and objectivity

Attracting candidates and ensuring reliability of evidence required that we maintain the anonymity of candidates. This was essential for fairness and to ensure that no one felt disadvantaged by having been prepared to participate in this pilot. Anyone failing to meet the standard which they were attempting might have been fearful about individual marks being fed back to the LSC or the regulatory bodies, particularly where they may have failed to achieve a pass at a level at which they were already practising. It was thus important to create a wall between the information about a candidate provided by the LSC, and the eventual identification number/name used for that person for the purpose of assessment.

Initial candidate information was submitted by candidates and collated by the LSC. This was passed in electronic form to the research team where it was managed securely. Candidates were given a Unique Identification Number (UIN) which appeared on all marking grids, with the addition of their name when it was an assessment when the assessor would meet the candidate. With portfolio assessments the only identifier sent to the assessor was the UIN. No other information was available to any assessor. Marks were assigned using the UIN and then cross-checked against their name before entry into the data base, to ensure that the correct result was assigned to each candidate. Before analysis the results were identified by number only but to this we now linked such other attributes as were important for the analysis of the data.

No candidate was notified of their mark. However it will be possible – were the pilot to find favour and become the template for a future scheme, and with the agreement of the LSC - for an assessment in the pilot to act as a passport to the relevant level.

3 Development of the pilot assessment scheme

3.1 Assessment development and pre-pilot

Once an assessment instrument had been identified as suitable for the pilot, criteria to enable all relevant competences to be marked in respect of that instrument were drawn up. The team of assessors is familiar with all the instruments used, and expert at matching a set of criteria to the competences requiring assessment.

Even experienced practitioners and assessors require criteria to ensure consistency of approach and the testing of the appropriate range of competencies. Our experience and the experience from peer review (work led by Professor Avrom Sherr of the Institute of Advanced Legal Studies in which Professors Cape and Moorhead have participated) suggests the benefits of a manageable number of criteria over more complex instruments. Large numbers of prescriptive criteria, designed for less experienced assessors as an audit mechanism (Transaction Criteria in Legal Aid being an example), become unwieldy in use.² When appropriately trained practitioner assessors are used, they can be provided with a slimmer more manageable document which is likely to be filled in completely. Streamlined criteria used by properly trained and monitored assessors ensure maximum consistency in reporting aspects of performance whilst allowing professionally experienced assessors properly to reflect experienced judgment.

Each assessment instrument produced by the team of assessors (so not the portfolio or written advocacy) was written by one or more members of the team and was subject to revision and scrutiny by at least two others to ensure it was free from unintentional ambiguity or elements which would cloud the ability of an assessor to obtain a true picture of a candidate's competence.

The case papers were supplemented with guidance to enable any assessor to identify the evidential and other points which the exercise had been designed to assess. The criteria used in the pilot were also devised and refined by the team of assessors, and tested as part of the process described below.

Each assessment instrument at Levels 1 and 2 was pre-piloted on a sample of advocates not participating in the main pilot. Those at Level 3 were dealt with in an abridged fashion, described later at 3.3.1. During the pre-pilot, live assessments were video recorded and the resultant performances subjected to joint scrutiny by all assessors, who discussed the relative merits/demerits of what they saw and decided on the standard to be reached. In pre-piloting of the witness handling, three pre-pilot performances were viewed by a team of practitioner assessors who all marked the performance "blind" without reference to one another or to the guidance, using the criteria created by the team.

² Complex assessment grids seen in the course of research for the pilot showed that such grids, far from enabling particularities to be noted, result in many aspects being left unmarked and the "General Comments" box being completely filled.

The aim of this initial exercise was to:

- ensure that the exercise allowed an assessor to make an evaluation based upon relevant competences;
- ensure that the exercise allowed an assessor to distinguish sufficiently between different levels of competence;
- check the completeness and accuracy of the guidance on points which should attract marks;
- ensure that the criteria and the layout of the mark grid allowed assessors accurately to record with sufficient particularity the aspects of a performance attracting marks, and those losing them;
- ensure that a basic level pass, and a fail at that level could be identified by reference to the use of the criteria and the attribution of marks;
- identify the characteristics of performances which should attract, or cause a reduction in, marks;
- ensure consistency of marking between assessors; and
- use the pre-pilot as an opportunity to train assessors in order to achieve greater consistency.

Each assessor reported and discussed the marks given. Inconsistencies of approach, such as a tendency not to use the full spread of marks, or a tendency to mark down an entire performance for inadequacy in one part only, were addressed. Points necessary to achieve success at each level were identified and included in revised guidance. The practicalities of using the criteria were discussed, for example, where on the marking grid particular aspects of performance resulting in the giving or deduction of marks should figure. The possibility of losing marks for the same failing under two or more parts of the marking grid was discussed and a common view reached as to when it was, and when it was not, appropriate. The guidance created by the designers of the case papers was distributed, and the assessors checked it contained all the relevant points.

3.2 Standards - cross level performance

Relative standards were discussed at the marking meeting. Given that the competences were drawn to operate across all levels, and that candidates identified several levels as applicable to their practice, it was felt important to seek to mirror this.

A pilot exercise which allowed testing of the same assessment instruments to assess candidates at different levels (see below), could do this. The marking meeting(s) of assessors discussed whether candidates who had undertaken an assessment at Level 1 had nevertheless reached Level 2 and vice versa in the live

assessments. Similar discussions were held about Level 3 candidates during the cross examination assessment.

If workable, such an exercise could address the problem that some advocates become confined to practice at a particular level and are thus unable to develop experience of cases at the higher level necessary to progress in their career. A system of accredited levels might inhibit such progress unless it contains within it a process for recognising the ability to move up levels and gain the necessary experience.

The creation of a rigorous assessment instrument, relying on a simulated piece of advocacy, which allows an advocate to prove themselves not only proficient at the standard at which they currently practise, but also has the possibility of recognising higher level performance would give opportunities to challenge unfair assumptions.

It would also mean that in any operational scheme there would be in place, as part of the testing process, an instrument specifically designed to assist in the kind of progression through the levels which must be achieved for the effective management of the Criminal Justice system.

With this in mind we adopted a case study from which witnesses for cross examination are drawn which would allow testing at Level 1 and Level 2. The dishonesty case (Taylor Annex H) used in the pilot allowed for the kind of detailed and specific questioning appropriate to assess Level 2 as well as Level 1. The pre-pilot established both this possibility, and where the divide between the levels lay.

3.3 Checking assessments in the pilot

3.3.1 Witness handling

On the first day of live assessment at Level 1 a sample of real “candidates” were called. Additionally three members of the team of advisers each undertook the regime of assessments, as if they were a candidate. Each of the live skills was marked by two assessors who awarded their own marks independently and then discussed those marks, in conjunction with the guidance, to reach a consensus. This process ensured standardisation of marking. One small change was made to the case study.

The original marking group consisted of eight assessors – but in practice, in order to achieve maximum consistency in an efficient way, only five of those eight were used for the Level 1 and 2 assessment team. Once an assessment day had taken place at each of the three centres, sufficient additional candidates had been assessed to merit a further meeting of that team, who were shown recordings of relevant assessments to discuss borderline cases, to ensure that there was still consistency and that no change in the standard being used was affecting the marks.

The case study was also scrutinised during the project by the Level 3 and 4 assessors and a senior judge. The judge did not see candidates perform but did

study the case and questioned the assessors and the witness. One of the Level 3 and 4 Q.C. assessors also conducted an assessment day of Level 1 and 2 assessments. Both that assessor and the judge expressed themselves satisfied with the appropriateness of the case study for the levels and its ability to allow distinction between those two levels.

For Level 3 cross examination a slimmed down version of the above process took place. Early meetings with senior members of the assessment team were devoted to design of the case study. A meeting was held to identify the nature of the exercise to be set, leading to a meeting with the witness – an expert. That exercise was then written by two members of the team and scrutinised by another two. The two QC assessors and the principal investigator met to discuss the case study and decide on tactics to adopt in rolling out this more complex case. Each of the QC assessors, the expert and the principal investigator wrote their own digest of points to expect to be made by the defence advocate and then by the prosecutor. All points were exchanged and discussed before compiling an appropriate set of notes for guidance for the assessors.

On each Level 3 assessment day two assessors were in attendance - though both might not be in the “courtroom” all the time. There was thus a ready opportunity to discuss the performances seen when they were fresh in the mind of the assessor, but a further meeting took place with all 3 assessors once all Level 3 assessments had taken place. On this occasion a number of the recorded performances were viewed and the mark to be attributed was agreed upon. This team took the view that 60% was an appropriate mark for that Level,

Having reviewed the performances which this exercise produced, the team of 2 Q.C assessors and the principal investigator reached a conclusion which could assist in designing a flexible operational scheme. They felt able to distinguish, quite apart from those not making the grade (achieving less than 60%), those whose performance was at the highest level – namely Level 4. Upon review of the recordings and of the spread of performances, the mark which divided this set of advocates from those at Level 3 was 80%. In an effort to utilise this distinction, the data shows those with the relevant mark as Level 4 for cross examination.

3.3.2 Multiple Choice Tests

Multiple Choice questions, targeted at assessing Levels 1 and 2 candidates without access to reference works, were written by two members of the team. The questions covered aspects of procedure and evidence, and also included appropriate questions of substantive criminal law. A bank of over 30 questions was produced before meetings were held to define the levels of the questions set (several were usable at both levels); to iron out any ambiguities and to redraft or reject any questions based on unfairly fine distinctions. Once this had been done the questions were put into two sets (Levels 1 and 2) and their subject matter considered. Where more than one question covered the same area, the better question was chosen. The set of questions was further reduced to the best spread of questions and three dual level questions selected which would appear in both

tests. The number of questions and the relevant correct answers to achieve a pass was decided upon.

The Multiple Choice Tests were checked by the team and trialled on the other assessors and members of the advisory group. Comments on clarity and correctness were considered and appropriate adjustments made. The Level 1 Test included questions on Law Procedure and Evidence, Level 2 on Procedure and Evidence only. (The test used are Annex F.) The very few Law questions the team felt appropriate at Level 1 were there because of the fact that leaving them there posed no unfair challenge. They should have been known to any criminal advocate without texts. Their inclusion was a valid test. However higher up the scale, the lack of a syllabus and the fact that at Level 2 the sort of non basic question appropriate would always mean the advocate would expect to look it up decided the team against any of the drafted law questions at this level. The questions were devised with one correct answer and three distracters. Marks were assigned to correct answers. Negative marks were not given to incorrect answers.

3.3.3 Portfolio and Written Advocacy

Portfolio and Written Advocacy criteria were drawn up by the academic assessors after consultation with the team of practitioner assessors. These were in part based on criteria with which the assessment team were familiar from the Criminal Litigation Accreditation Scheme, but there were significant modifications to account for the assessment of trials and high level work. The pilot did not allow the inclusion in a portfolio of a case where the candidate was not the lead advocate. To do so risked giving credit to a candidate for decisions, thought processes, advice, tactics or legal argument which has not stemmed from that advocate and would make evaluation of the assessment scheme less reliable. Once in use, certain criteria were modified during the pilot to take account of difficulties faced by participants in complying with all requirements or because it was found that in practice one criterion was sufficiently covered by another.

Once each set of portfolios for a level had been received they were divided between two assessors. Each pair of assessors received copies of two portfolios also being marked by the other. Each blind marked these, and then exchanged and discussed their marks to consider issues of consistency in their use of the criteria.

Once the benchmarks were thus set the rest of the portfolios were marked before a meeting took place, at which were present the principal investigator, Level 1 and 2 assessors and one of the Level 3 assessors, who had meanwhile been marking the higher level portfolios. That meeting gave an opportunity to iron out any problem case, but also to report back on the use of the criteria. As a result of this meeting and another such, the criteria were modified slightly to the form in which they are to be found at Annex D and that modified form was used in reaching the final marks given to the portfolios marked.

3.3.4 Judicial Evaluation

The judicial evaluation form was designed by the assessment team – one of the QC assessors and the principal investigator. A short set of explanatory notes was also developed. It was pre-piloted in Southwark Crown Court where it was reported by the liaising member of the Assessment Workstream Group that the judges who tried it found it relatively easy to use. An amendment was made after this trial and the form was then introduced to the main pilot test centre courts.

We held meetings at three courts to explain the pilot and our aim in enlisting their support, also explaining where necessary the use of the form. It was not possible to train the judges in use of the form or to engage in a process for ensuring (through pre-pilot work with them) that they reached common standards. Indeed a judge would have no possibility of comparing the standards of an advocate before them in a real case with the standard of an advocate before another judge in another real case. One possibility to ensure judicial consistency in the use of such criteria would be to monitor the marks they give to advocates over a larger sample and see whether some judges are tougher (or easier) in their assessments than others and explore the reasons for that. That was not possible in the context of the number of judicial evaluations likely to be, or actually, carried out in this pilot.

To undertake any training in the use of the form would have involved significant resources and involvement by the Judicial Studies Board. It could not, in any event, have been achieved within the time frame of the QAA project. In the explanatory notes any judge desiring further explanation was invited to telephone or e mail the principal investigator. In the event, two judges did so. Informal feedback from judges about the content of the form was wholly positive.

3.4 Assessment criteria

The various sets of assessment criteria used for the pilot appear at Annex D. Their specific attributes are more particularly discussed in the report under the general explanation of the regime of assessments for that level, and also where appropriate under operational considerations.

All marking criteria have in mind the Competency Framework agreed by the QAA Reference Group but each set of criteria uses only such elements of that Framework as are appropriate for the relevant Level(s) and instrument. The criteria have to recognise the fact that no assessment instrument can give direct evidence of all of the framework competences.

In devising the criteria the team were concerned about the fact that some of the Competences we had inherited are expressed in a way which rewards only perfection e.g. 'asks **only** relevant questions' or 'is courteous at **all** times'. That wording makes them incapable of gradation. We did not believe that this had been the intention of the group which devised them and therefore such competences were always read – where they were applicable – as if they intended partial achievement could gain some marks.

3.5 Finalising data

Data from the assessment mechanisms was checked by the principal investigator for arithmetical and internal consistency and then entered onto a spreadsheet. Gaps in data were followed up with assessors and candidates where appropriate. Certain candidate information was coded into more generalised categories to aid analysis. For example:

- the variety of degrees was reduced to Law; Law and another; Non Law or None.
- where an advocate had taken the BVC (graded Outstanding, Very Competent, Competent) their grade was inserted when known. Where the course giving a barrister their professional qualification predates the BVC it was entered as a Pass and thus indicates **no** grading.
- where an advocate had taken the LPC (graded Distinction, Commendation, Pass), their grade was inserted when known. Where the course giving a solicitor their professional qualification predates the LPC it was entered as LSF or Part 2 and thus indicates **no** grading.

The professional group to which a candidate belonged was recorded in order to better analyse data by noting any differences by grouping and seek any data to explain these as well as to feed into any recommendations which might be able to be made about passporting or exemption.

4 Cohort

4.1 Background

When the pilot was first conceived, it was felt that for operational efficiency, and in order to have the maximum opportunity to marry up a judge's evaluation of a candidate with that made by the use of assessment instruments, the candidates should be drawn from a limited number of court centres. Those centres were Inner London, Birmingham, Winchester, Cardiff and Newport

In order for both the bench and the advocates in those centres to understand what might be required of them in the pilot, several events were held where those most concerned had an opportunity to obtain information from members of the research team and the LSC team, upon which to base their decision about participation. The recruitment of candidates for the pilot was delayed by this process, but even after those meetings volunteers did not come forward in sufficient numbers.

One way to achieve greater participation was to uncouple the assessment structures from the designated court centres. In this way we were more likely to achieve the spread of experience to test the assumptions that we so keenly desired to do.

The “uncoupling” did occur, linked with the possibility of Judicial Evaluation from those courts too. The numbers of volunteers increased markedly, though too late in the day to enable full advantage to be taken of this increased pool of participants – particularly when some were practising at a great distance from the assessment centres in Cardiff, Birmingham and London.

In an effort to increase the opportunity to attend, two additional assessment days were set up at the request of particular sets or firms in Nottingham and Newcastle, areas not designated as assessment centres.

4.2 Cohort make up - general

Candidates who volunteered to take part in the pilot had several potential reasons to put themselves forward. Some candidates indicated a desire to ensure that in future clients would be represented by the best people and the public purse used efficiently; others that if some quality assurance scheme was to be implemented, they would like to have a practice at it and maybe thus also shape it. Finally there was the hope held by some that in the future they would gain exemptions in an operational scheme by successful participation in the pilot. No CPD points were able to be offered during the pilot to encourage participation.³

It is important to recognise that those taking the pilot assessments voluntarily may not be typical of the professions as a whole. In particular, we do not know whether the levels of competence that the volunteer candidates demonstrated would be

³ Since the time for volunteering closed, the professional bodies have allowed the volunteers' contribution to be recognised by such an award.

higher, lower or similar to those to be expected under a mandatory accreditation scheme. Whilst one might expect that those participating as volunteers would be more confident than average of their competence, the fact that many candidates might use the exercise as a way of gearing up for the scheme and testing themselves at the level they aspired to, might mean that their results would be poorer than would be expected should a scheme be rolled out beyond the pilot. The results reported below need to be read in that light.

We are also aware that the judiciary played a role in participation. Some judges championed, and some actively discouraged, participation. Communications from candidates about this led us to believe that this impacted most significantly on the bar, but also on other Crown Court advocates appearing regularly before such judges.

4.3 The cohort

4.3.1 Who was assessed?

As noted above, of 227 potential candidates, 101 were assessed and 98 produced data in time for the evaluation using one or more of the QAA assessment mechanisms. Of these 10 were also evaluated by a judge. 12 further candidates were evaluated by judges but not by other assessment mechanisms. Only three candidates were evaluated by two judges. To distinguish between the two assessment mechanisms we refer to 'assessment(s)' for the assessment instruments designed and implemented by the Cardiff teams, and 'evaluation' for the judicial feedback provided on candidates' performances in court.

Table 2: Assessed

		Frequency	Percent
Valid	Assessed	88	38.8
	Assessed & Judicially Evaluated	10	4.4
	Judicially evaluated only	12	5.3
	Unable to participate	97	42.7
	Withdrawn	20	8.8
	Total	227	100.0

4.3.2 Who was not assessed?

Barristers, CPS and Filex candidates were more unlikely to participate in assessments than solicitor candidates (Table 3) but this is partly explained by the spread of professions at different levels.

Table 3: Participation by professional grouping

	Barrister	CPS	FILEX	Pupil	Solicitor	Total
Assessed	26%	30%	38%	50%	52%	88
Assessed+Judicially evaluated	9%	4%	0%	0%	1%	10
Judicially evaluated only	11%	4%	0%	0%	1%	12
Unable	49%	43%	38%	17%	39%	97
Withdrawn	5%	17%	25%	33%	7%	20
Total	92	23	8	6	98	227

The way in which rates of attrition operate across the levels is shown in Table 4. The rate is generally high in Levels 3 and 4 and this is where a large proportion of the barristers were indicating they would wish to be assessed. Even so, solicitors were less likely to withdraw or be unable to be assessed at Level 2 than Barristers who were assessable at Level 2. The differences were not however statistically significant.

Table 4: Attrition by level and professional grouping (percentage either withdrawing or unable to sit assessments)

Level	Barrister	CPS	FILEX	Pupil	Solicitor
1	64%	40%	38%	50%	53%
N	11	5	8	6	38
2	47%	38%	0%	0%	70%
N	19	13			40
3	28%	0%	0%	0%	29%
N	36	2			14
4	23%	33%	0%	0%	0%
N	26	3			6

The following analysis concentrates on the 98 candidates assessed using the assessment mechanisms and the 22 candidates for whom there are judicial evaluations.

The pool of candidates was predominantly male (63 out of 98 (64%)). The average (mean) age of candidates was 41 years. The nature of their degree qualification is set out in Table 5. Additionally, 26 were identified as having done the CPE (6 of those also said they had done a Law degree).

Table 5: Candidates' Degree

		Frequency	Percent
Valid	Law	65	66.3
	Law and Other	7	7.1
	Non-Law	19	19.4
	None	6	6.1
	Missing	1	1.0
	Total	98	100.0

The average length of PQE or years call was 15 years. Averages were similar for barristers, solicitors and CPS advocates. However, of those presenting for assessment at Level 1, there was a significant difference between the average length of experience in criminal work for solicitors (17.4 years) compared to barristers (1.9 years).

38 (73%) of solicitor candidates had Higher Rights of Audience: 30 of these via the exemption route, 5 through accreditation and 3 through the developmental route. 48 (92%) of the solicitors were also Duty Solicitors.

5 Results from the assessment mechanisms

As noted in section 2.4, candidates took different combinations of assessments depending on the Level at which they were being assessed. This section of the report sets out the results from each assessment and then compares the results of different assessment mechanisms. This enables some empirical analysis of the extent to which the assessment mechanisms relate to each other or appear to be assessing different elements of competence. The results of the judicial evaluation are then considered in their own right (Section 6) and in comparison with the other assessment mechanisms (to the extent that this is possible given the low number of judicial evaluations).

As already noted above, it needs to be emphasised that the sample of candidates being assessed here may not be typical of the general population of advocates in the professions.

Similarly across all levels, but particularly Level 3 and Level 4, the numbers participating in the assessments is relatively small. This coupled with the atypical nature of the candidate sample means the results should be interpreted with significant levels of caution.

5.1 Level 1

The case study used for the interview/conference and submission assessment appears as Annex G (the case of Cooper). The defence candidates interviewed the client (played by an actor) in respect of a difficult s.47 assault occasioning actual bodily harm, requiring some thoughtful questioning in order to advise on plea and, thereafter, venue. There were also bail issues. These matters then formed the basis of the simulated Court “submissions”. The CPS candidates were given a different take on the same case and had to explain to the victim witness attending for trial a recent development which could lead to a plea. Their submission required the candidate to introduce the case for what was now a sentencing hearing – taking account of some information gleaned in the interview.

When deciding on a case study to be used for the witness handling assessment at Level 1 the team chose an either way matter which could have found its way into the Crown Court by election or remained in the magistrates’ court. The case study is that of Taylor and appears as Annex H. The case was one of theft and the exercise provided as a defence exercise in cross examination for exploration of up to seven relevant issues. For CPS candidates cross examining the defendant there were slightly fewer issues but the questioning needed to be more challenging on some.

Performances which lacked depth or insight (and generally these were the ones taking significantly less than the 20 minutes allowed) would fail, as would those not dealing with essential points. The cross examination could last up to 20 minutes but a candidate would not fail for not completing the exercise in the time, except if the reason for this was irrelevant or imprecise questioning. Their written plans were

available for assessors to check on the issues they had not covered within the time, but in the event all substantially managed the exercise in the time given.

5.1.1 The results of Level 1 assessments

Candidates undertaking Level 1 assessments were required to sit an interview (conference), a submission (both interview and submission were based on the same case-study), a cross examination assessment, and a multiple choice assessment. They were also required to submit a written advocacy assessment and a portfolio, although relatively few candidates did.

Failure rates on these assessments were generally low. This may be consistent with the volunteer status of the candidates. There could be no tendency to seek validation at a higher level (as this was the lowest level) and on balance one might expect only the more confident Level 1 candidates to come forward. Of 35 candidates who took Level 1 assessments:

- 4 (11%) failed the interview
- 5 (14%) failed the submission
- 2 (6%) failed the cross examination assessment. They both passed their interview. One had also failed on their submission.
- 3 (9%) who failed the interview reached Level 1 on cross examination and the other interview fail reached Level 2. 4 of the submission fails reached Level 1 (one did not).
- 8 (23%) Level 1 candidates failed the MCT
- 2 out of 7 (29%) of those submitting portfolios at Level 1 failed. Both fails passed the interview assessment. 1 candidate passed their portfolio but failed the interview.
- Only 4 candidates submitted written advocacy. There was one technical fail where the candidate submitted a piece of advocacy from a civil matter not eligible for assessment.

In terms of similarity across the assessment mechanisms:

- 3 of the 4 who failed the interview also failed the submission (3 of the 5 who failed the submission also failed the interview).
- Of the 2 who failed the cross examination assessment, both passed their interview. One failed on their submission.
- 3 who failed the interview reached Level 1 on cross examination and the other interview fail reached Level 2. This person failed the interview for the very characteristic which made their cross examination so effective – and it was to do with their rapport with their client. 4 of the submission fails reached Level 1 (one did not).

- 8 Level 1 candidates failed the MCT. 2 of the 5 who failed the submission also failed the MCT. 6 of the 8 who failed the MCT passed the submission. 2 of 4 interview fails also failed the MCT (6 MCT fails passed the interview). 1 of the MCT fails also failed cross examination. 4 passed at level one and 3 got to Level 2.
- 2 out of 7 people failed on their portfolios. Both fails passed at interview. 1 candidate passed portfolio but failed at interview.
- Only 4 candidates submitted written advocacy. There was one technical fail where the candidate submitted an inappropriate case for assessment.

Level 1 candidates assessed as performing at Level 2

In cross examination it was possible for the candidates being assessed at Level 1 to be assessed as performing at the next level. 12 out of 35 (34%) Level 1 candidates reached Level 2 on cross examination (but one of those failed their interview).

We looked at this data by a number of candidate characteristics.

Table 6: Cross examination level by professional group

Professional Grouping		Cross examination Level 1 Candidates			
		Fail	Level 1	Level 2	N
Barrister	N	0	5	2	7
	%	.0%	71.4%	28.6%	100.0%
CPS	N	0	1	1	2
	%	.0%	50.0%	50.0%	100.0%
FILEX	N	0	3	0	3
	%	.0%	100.0%	.0%	100.0%
Pupil	N	1	1	1	3
	%	33.3%	33.3%	33.3%	100.0%
Solicitor	N	1	11	8	20
	%	5.0%	55.0%	40.0%	100.0%
	N	2	21	12	35
	%	5.7%	60.0%	34.3%	100.0%

Although proportionately more solicitors did well, the difference was not statistically significant⁴.

We compared those solicitors at this level who had higher rights with those who did not. Those who did not, tended to get better scores on the cross examination assessment, though again this was not statistically significant.⁵ All of the solicitors with higher rights of audience in Level 1 had gone through the exemption route.

4 Kolmogorov-Smirnov Z, p = .26

5 Kolmogorov-Smirnov Z, p = .375

Any solicitor with higher rights taking the assessment at Level 1 had requested assessment at this level only. Those of them to whom we spoke about the correct level at which to assess them explained that they did not exercise their higher rights in trials.

Table 7: HRA by cross examination result (Level 1)

			Cross examination			
			0	1	2	N
HRA	No	Number	0	7	6	13
		%	.0%	53.8%	46.2%	100.0%
	Yes	Number	1	4	2	7
		% within HRA	14.3%	57.1%	28.6%	100.0%
Total		Number	1	11	8	20
		%	5.0%	55.0%	40.0%	100.0%

Although those Level 1 candidates performing at Level 2 had slightly higher average experience (in terms years of criminal PQE) than those performing at Level 1, the difference was very small (less than one year on the averages) and not significant).⁶

There were no observable trends or significant differences in the distribution of cross examination scores for different degree types or gender, LPC or BVC grade.

These results do not suggest any empirical basis for suggesting passporting requirements for Level 1.

5.1.2 A comparison of the assessment mechanisms - Level 1

We have analysed the results from the different assessment mechanisms to see what level of agreement, if any, there is between them. Where there is agreement between the mechanisms there is greater reason for believing that the assessment mechanisms are looking at the same or related elements of competence.

Depending on the strength of that agreement and the rationale for including each test, this may lead to a view that only one or other of the tests should be used (or some less onerous combination of the tests). Where the tests do not give a similar result, it may indicate that one or other of the tests is unreliable (again this depends on other supporting evidence and in particular doubts about the rationale for a particular test) or that the tests are measuring different aspects of professional competence. Judgments in this regard are complicated by the relatively low take up for the tests, the generally high pass rates for the tests at Level 1 and the potential for this cohort to be skewed towards good performance.

We investigated the levels of agreement between different assessment mechanisms by looking for correlations between them. Correlations examine the results for each candidate on a pair of tests and assess whether, for all candidates doing both assessments, those that generally do better on one test also generally

⁶ T-test, $p = .877$.

do better on the other test with which it is being compared. In the main, the variables were looked at on a pass-fail basis.⁷ Correlation coefficients vary between -1 and 1, with a coefficient value that of 1 indicating total agreement and -1 indicating total disagreement between the measures.

In making such comparisons the report identifies any relationships which are 'statistically significant'. The report uses the conventional approach of assessing significance. We only identify relationships as significant where the probability of falsely identifying a relationship when there is not in fact a relationship is less than 1 in 20 (.05). That is, when a test is used to identify a similarity between one assessment mechanism and another, the probability of that similarity occurring by chance is less than 1 in 20 ($p < .05$). The convention tends towards the prevention of false positives (that is it protects against identifying a relationship between two variables which does not in fact exist). As a result, the convention is more prone to the identification of false negatives (suggesting there is no relationship when in fact there is). To counter this risk in part, we also report, the probabilities as being 'near significance' where the probability of a false positive is less than 1 in 10 ($p < .1$). This also gives some greater depth to the data analysis given the small number of assessments carried out at each level.

Relatively few Level 1 candidates failed their Level 1 assessments, making apparent relationships between the different assessment mechanisms less likely. The results of correlation tests for Level 1 assessment mechanisms are summarised in the following table. Correlation coefficients were generally low, save for the comparison of pass-fails in interviews and the submission assessment, where a higher level of agreement was found which was near to (but did not reach) conventional levels of statistical significance. The correlation coefficients tended to be positive, suggesting some (albeit not statistically significant) agreement, save notably for portfolio assessments where there was a measure of disagreement between portfolio and other scores (where there were in fact only 7 assessments).⁸

⁷ Correlations between two pass-fail assessments were conducted using Kendall's tau-b. Generally scales which went beyond pass-fail were reduced to a pass fail level initially to compare with other mechanisms. Where the assessment produced marks which could be treated as a scale (such as the level achieved in a cross examination assessment) Pearson correlations statistics were examined and are reported if significant and different from the pass-fail comparisons conducted as the main focus of the analysis.

⁸ The interview cross examination correlation coefficient is negative but close to zero.

Table 8: Level 1 mechanisms compared⁹

		Kendall's tau-b	p
Interview	Submission	0.623	0.055
Interview	Cross examination	-0.088	0.210
Interview	MCT	0.232	0.288
Submission	MCT	0.167	0.407
Cross examination	Submission	0.251	0.382
Cross examination	MCT	0.159	0.478
Portfolio	Interview	-0.258	0.299
Portfolio	Submission	-0.258	0.299
Portfolio	MCT	-0.091	0.811

One further set of comparisons was carried out. Within the interview, the “advising the client” criterion was felt by the assessors to be the most important element of the assessment; likewise, “persuasiveness” in the submission and “achieving the objective” in cross examination. These elements were scored on a scale and so the scores were examined to see if there was a correlation between them. Advising the client in the interview submission was significantly correlated with the persuasiveness of the submission¹⁰ but achieving the client’s objectives in cross examination was not significantly correlated with the persuasiveness of the submission¹¹ or advising the client in the interview.¹²

These results demonstrate therefore, that there was a reasonably strong overlap between the interview and the submission assessments, which is reinforced when one looks at the key elements within those assessments. That is not to say that there is no variation between the two types of assessment, but if other evidence points in the same direction it might justify relying on one or other of the assessments. Beyond that, with this set of candidates, there was little clear agreement between the assessment mechanisms as to which candidates were competent or not competent. This suggests either problems with the assessment mechanisms themselves (something which the pre-piloting and scrutiny which they underwent should allow us to discount) or that the mechanisms were assessing somewhat different aspects of competence.

The results suggest so far, albeit on modest numbers of candidates, that an approach that assessed candidates across a range of assessment mechanisms, would provide a sufficiently wide and reliable accreditation mechanism. However,

⁹ This comparison only considers pairs of assessment when some candidates failed both assessments. This is because where all candidates taking one assessment passed the other assessment there is no variation to be measured by way of correlation.

¹⁰ Kendall's tau_b .435, p = .002

¹¹ Kendall's tau_b -.156, p = .295

¹² Kendall's tau_b .173, p = .244

21 out of 35 (60% of) Level 1 candidates failed at least one assessment, suggesting that failure on any one mechanism may be unduly harsh. Of the four compulsory assessment carried out at Level 1 and using the pass rate on the must-knows on the MCT test rather than the overall score on the MCT (which represents the sterner test) only 5 (14%) candidates failed 2 of the compulsory assessments (and only one other candidate failed one compulsory test and a portfolio submission). The results for these candidates are summarised in Table 9.

Table 9: Summary of Level 1 failures

Inter-view	Sub-mission	Cross-exam	MCTs		Portfolio	Written Advocacy
			Overall	Must-knows		
Fail	Fail	Pass	Fail	Fail	.	.
Fail	Fail	Pass	Fail	Fail	Pass	.
Fail	Fail	Pass	Pass	Pass	.	.
Pass	Fail	Fail	Pass	Pass	.	.
Pass	Pass	Fail	Fail	Fail	.	.
Pass	Pass	Pass	Fail	Fail	Fail	.
Fail	Pass	Pass	Pass	Pass	.	.
Pass	Fail	Pass	Pass	Pass	.	.
Pass	Pass	Pass	Fail	Fail	.	.
Pass	Pass	Pass	Fail	Fail	Pass	.
Pass	Pass	Pass	Fail	Fail	Pass	Pass
Pass	Pass	Pass	Pass	Fail	.	.
Pass	Pass	Pass	Pass	Fail	.	.
Pass	Pass	Pass	Pass	Fail	.	.
Pass	Pass	Pass	Pass	Fail	.	.
Pass	Pass	Pass	Pass	Fail	.	.
Pass	Pass	Pass	Pass	Fail	.	.
Pass	Pass	Pass	Pass	Fail	.	.
Pass	Pass	Pass	Pass	Fail	Pass	Pass
Pass	Pass	Pass	Fail	Pass	.	.
Pass	Pass	Pass	Pass	Pass	Fail	.

5.2 Level 2

Level 2 candidates had a different mix of assessment mechanisms to sit: a cross examination lasting 20 minutes; an examination in chief of up to 15 minutes; a multiple choice test; and, a portfolio and/or a written advocacy assessment. The witness handling derived from the case of Taylor referred to above. There were generally lower pass rates on some of the assessments (the multiple choice test, cross examination and examination in chief in particular) at this level. This may be explained by the nature of the candidates (more of whom may have been seeking to demonstrate performance at a level higher than the one at which they normally practised). Failure in either witness handling exercise included, for a Level 2 candidate, reaching a mark only sufficient for Level 1.

5.2.1 The results of Level 2 assessments

20 out of 39 (51) of those taking the cross examination assessment at Level 2 passed the assessment. 22 out of 39 (59%) passed the examination in chief.

The results of the two assessments are compared in Table 10.

Table 10: Cross examination compared with examination in chief

Level			Examination in Chief Level 2		
			Fail	Pass	Total
2	Cross examination Level 2	Fail	15	4	19
		Pass	2	18	20
	Total		17	22	39

The scores on these two assessments are significantly correlated suggesting a strong relationship between the two assessment mechanisms.¹³ It is worth noting also that of the 6 that failed one of the witness handling tests, 4 failed another assessment (3 failed the multiple choice test and 1 the written advocacy assessment).

Given the level of failure under this assessment it is worth outlining the reasons candidates failed. The main reason for failing at Level 2 was a lack of thoroughness in selecting the issues that were ripe for cross examination. Such candidates failed to properly address the “theory” of the prosecution or defence case and failed to test a number of assumptions made by the key prosecution witness. The cross examination needed to be properly planned to draw out inconsistencies and weaknesses that were evident from other witness statements and for many candidates this was not done. It was not a requirement that each and every issue in the witness’s statement needed to be tested in order to achieve Level 2.

¹³ Kendall's tau-b, 0.695, p =0.000.

The failing Level 2 candidates failed to demonstrate any level of sophistication in their cross examination. There was a very obvious issue to be tested and all candidates made some effort to test this in cross examination. Those candidates that demonstrated thorough probing of the key issues in the case were able to attain Level 2 marks overall, even if they missed some of the other issues that were appropriate for cross examination. The marking scheme allowed for such candidates to pass provided that they presented well, with a strong questioning technique and good structure.

Furthermore, the assessors noticed that those candidates lacking a sophisticated approach to cross examination tended also to struggle with the structure of the questioning. It was often evident that the candidates that had failed to properly address their minds to a case theory were haphazard in their questioning structure. As a result the questioning would often lack coherence and overall such candidates (if for the defence) both failed to undermine the prosecution case and properly advance the defendant's.

20 out of 39 (51%) passed the Multiple Choice Test. There was no correlation between pass:fail on cross examination and Multiple Choice Test pass:fail¹⁴ as can be seen clearly in the following table.

Table 11: Cross examination Level 2 compared with Multiple Choice Test

Level			Multiple-choice test		Total
			Fail	Pass	
2	Cross examination Level 2	Fail	9	10	19
		Pass	10	10	20
	Total		19	20	39

There was understandably some uneasiness among the candidates about sitting a Multiple Choice Test. They had not been given a syllabus or guidance as to any revision to undertake. They were not to be allowed to refer to texts. In setting the tests the assessors were mindful of the fact that this was an unseen closed book assessment and our aim was to pose questions the answers to which in the view of the assessors should be known by competent advocates. Nonetheless in exit feedback it was the Multiple Choice Test which a number of the candidates were most fearful would have uncovered gaps in their knowledge. A regular complaint was that they would have looked up the answers to these questions.

The Multiple Choice Test at Level 2 contained no substantive law questions, and asked more Evidence (8) than Procedure (5) questions because at this Level the advocate will be increasingly likely to be able to look up substantive law in advance. Bearing in mind the feedback about the tendency of practitioners to look matters up, the team reconsidered the Multiple Choice Test at Level 1 and Level 2 and identified 4 questions at Level 1 and 5 questions at Level 2 which the research

¹⁴ Kendall's tau-b -.026, p = .869

team and the assessors felt that candidates should know without having to look them up, and where not knowing them could seriously compromise a trial or the client's case. We reviewed the marks for the Multiple Choice Test using the marks for those questions alone to see whether a pass (4 out of 5 at Level 2) on these "must-know" questions bore any significant correlation with other important aspects of performance. A pass-fail on the must-know scores in the Multiple Choice Test was modestly but significantly correlated with the pass-fail on the overall Multiple Choice Test score.¹⁵ Otherwise, the results were very similar for a comparison of Multiple Choice Tests and examination-in-chief: they failed to show any correlation.¹⁶

The score for the 'Achieving the objective' criterion in cross examination showed no statistically significant correlation with the scores on the Must-know MCT questions at this Level, or at Level 1.

21 out of 39 (54%) passed the must-know questions (fails scored less than 4 out of 5 of these correctly). Again, there was no correlation between this result and the pass-fail in cross examination¹⁷ or examination in chief.¹⁸

19 Level 2 candidates submitted portfolios and all passed. 20 submitted written advocacy assessments. There were two 2 fails and one technical fail. This fail, like the one at Level 1, resulted from the submission of an inappropriate piece of work, and thus did not reflect on the standard of the candidate so much as their uncertainty about the process.¹⁹ There was a near significant correlation between written advocacy and examination in chief based on the two candidates failing in written advocacy also failing examination in chief.²⁰

Table 12: Written advocacy compared with examination in chief

			Examination in Chief Level 2		Total
			Fail	Pass	
2	Written Advocacy	Fail	2	0	2
		Pass	4	11	15
	Total		6	11	17

¹⁵ Kendall's tau-b 0.49, $p = 0.003$.

¹⁶ Kendall's tau-b -.029, $p = .855$. We have not included a table given the similarity with Table 11. An analysis comparing the most important criterion in the cross examination (achieving the objective) with Multiple-Choice Test scores did not suggest any relationships either.

¹⁷ Kendall's tau-b, .024, $p = .882$.

¹⁸ Kendall's tau-b, .120, $p = .453$.

¹⁹ Which is treated as missing data in the analysis.

²⁰ Kendall's tau-b .494, $p = .094$

These results suggest therefore that there was a strong overlap between the assessments provided by cross examination and examination in chief. That is not to say that there is no variation between the two types of assessment, but if other evidence points in the same direction it might justify relying on one or other of the assessments. There is also some suggestion of a relationship between the written advocacy and examination in chief assessments, although this is based on a relationship between a very small number of fails.

5.2.2 Cross-over questions in the Multiple Choice Test

There were three questions which were answered both by Level 1 and Level 2 candidates. Although limited in number they give some indication of whether the Level 2 candidates did better than Level 1 on those questions.

Table 13: Scores on the MCT crossover questions by Level of Assessment

			Assessment Level		
			1	2	
MCT cross-over scores	0	Number	2	2	4
		%	5.7%	5.1%	5.4%
	1	Number	17	13	30
		%	48.6%	33.3%	40.5%
	2	Number	12	16	28
		%	34.3%	41.0%	37.8%
	3	Number	4	8	12
		%	11.4%	20.5%	16.2%
Total		Number	35	39	74

5.2.3 Considering a range of assessment mechanisms (Level 2)

A sense of how candidates performed across a range of assessments can be seen in Table 14. 30 out of 42 (71%) Level 2 candidates failed 1 or more of the main tests (multiple choice, cross examination and examination in chief), although often this was the cross examination and examination in chief (emphasising the similarity of competences assessed). 17 (40%) failed the MCT and at least one of either cross examination or examination in chief.

Table 14: Summary of candidates with at least one fail (Level 2)

MCT	Cross examination Level 2	Examination in Chief Level 2	Portfolio	Written Advocacy
Fail	Fail	Fail	Pass	Pass
Fail	Fail	Fail
Fail	Fail	Fail
Fail	Fail	Fail	Pass	Pass
Fail	Fail	Fail	Pass	Pass
Fail	Fail	Fail
Fail	Fail	Fail	Pass	Pass
Fail	Pass	Fail
Pass	Fail	Fail
Pass	Fail	Fail	Pass	Pass
Pass	Fail	Fail
Fail	Fail	Pass
Pass	Fail	Fail
Pass	Fail	Fail
Fail	Fail	Pass	Pass	Pass
Pass	Fail	Fail
Pass	Fail	Fail	Pass	Pass
Pass	Fail	Fail
Fail	Pass	Pass
Fail	Pass	Pass
Fail	Pass	Pass	Pass	Pass
Fail	Pass	Pass
Fail	Pass	Pass
Fail	Pass	Pass	Pass	Pass
Fail	Pass	Pass	Pass	Fail
Fail	Pass	Pass	Pass	Pass
Fail	Pass	Pass
Pass	Fail	Pass
Pass	Fail	Pass	Pass	Pass
Pass	Pass	Fail
...	Pass	Fail

5.3 Level 3

Candidates at Level 3 were assessed in cross examination and were also invited to submit written work in the form of a portfolio case and piece of written advocacy. The difficulty of creating a sufficiently complex complete set of papers for high level performances in cross examination at Level 3 – and possibly 4 - meant that the assessment was based on a part only of a case, but one which required the advocate to show competences at a higher level.

The initial assessment team meeting considered a variety of offences capable of being tried at Level 3. We decided that, in order readily to have access to a witness who would consistently present the same challenge for each candidate, an expert witness would best serve the purpose. In this way the exercise could also be intellectually challenging and pose tactical considerations.

Using an expert also allowed the assessment to focus on an element of a more complex trial in a manageable way. It was possible to ensure that the assessment contained sufficient information to test the advocate's grasp of the importance of the evidence within the context of a complex trial; assessing the ability of the advocate to deal with 'red herrings', and requiring tactical decisions both before the assessment and on the day.

Consideration was given to what kind of real expert would be available regularly at an economic rate for the assessments, where there would not be the need to provide access to a candidate's own expert. We needed one where the nature of the expertise was within a field which readily allowed a lawyer to research themselves, and with limited prompting by way of paperwork from us. Such self-directed research and analysis was consistent, in the view of the research team and the assessors, with Level 3 advocacy.

The case (see Annex J) was one relating to fraud from overcharging by a solicitor. The expert to be cross examined was a costs draftsman. The exercise was to last 40 minutes. This did not reflect the length of time that cross examination in a real case would really need – but was a time within which the assessors felt they would be able to judge the competence of the advocate. That belief was supported by the performances. The exercise was received by the advocates at least 3 weeks before their assessment appointments. The paperwork directed them to read it early and to apply for any additional documents or information which it appeared upon consideration would assist²¹.

On the day, the candidates were given 40 minutes before their cross examination within which to do final preparation. Those who had not sought some crucial documents were given them then. Other useful documents were however disclosed only upon prior request, on the basis that the failure to request such documents was an important indicator of the advocate's competence.²² Those embarking on questioning the witness without this information lost marks. Those who had sight of the additional documents only 40 minutes before were put in a challenging situation. As they had been notified in their paperwork, the advocate's notes or plan for cross examination were collected in after the exercise and contributed to the way in which the performance was scored.

²¹ The candidates were told that not all documents which they might like would have been prepared for this exercise, but that such as were available would be sent to them upon inquiry and an explanation given as regarded others. Those who sought information early would therefore have the benefit of being better prepared.

²² For example the expert's c.v. was available only by prior request.

5.3.1 The results of Level 3 assessments

10 out of 13 (77%) passed cross examination. 4 (31% of Level 3 candidates taking the assessment) reached Level 4, that is achieved a mark of 80 or more for their cross examination. 8 out of 10 candidates passed the portfolio. There was no significant correlation with the pass-fail in cross examination.²³ Nor did written advocacy correlate with cross examination.²⁴

Written advocacy and the portfolio had a correlation which was near significant, although only 11 candidates did both, with the written advocacy test being harder to pass.²⁵

Table 15: Portfolio and Written Advocacy compared (Level 3)

Level			Portfolio		N
			Fail	Pass	
3	Written Advocacy	Fail	2	2	4
		Pass	0	7	7
	Total		2	9	11

Most advocates had chosen their portfolio cases well, and a number of candidates had produced pieces of work relating cases between which, and those submitted at Level 4, the assessors found no distinction. The requirement of leading, which is the additional facet at Level 4, was either present in the advocate's team in these portfolio cases, or implicit in the utilisation of team members who would not be on their feet.

²³ Kendall's tau-b, .375, $p = .364$.

²⁴ Kendall's tau-b, .102, $p = .753$

²⁵ Kendall's tau-b .624, $p = .058$.

5.3.2 A comparison of the assessment mechanisms - Level 3

Performance across the range of assessments can be seen in Table 16. Two out of fourteen candidates failed on two or more mechanisms.

Table 16: Performance across Level 3 mechanisms

Portfolio	Written Advocacy	Cross examination
.	.	Higher Pass (L4)
Pass	Pass	Pass
Pass	Pass	Higher Pass (L4)
Pass	Pass	.
Pass	Fail	Pass
.	.	Higher Pass (L4)
.	.	Fail
Pass	Pass	Higher Pass (L4)
Pass	Fail	Pass
Fail	Fail	Pass
Pass	Pass	Pass
Pass	Pass	Fail
Pass	Pass	Pass
Fail	Fail	Fail

5.4 Level 4

At the time when the diet of assessments appropriate to a candidate at Level 4 was being considered, no Level 3 candidates had been tested. As a result the team had reservations about attempting to develop a live assessment which was sufficiently testing to enable an assessment of Level 4 whilst also being manageable in terms of the size of the set of papers that needed to be prepared and considered by candidates. As a result the assessment instruments chosen for Level 4 were the portfolio and written advocacy.

5.4.1 The results of Level 4 assessments

For reasons discussed above, the number of candidates at Level 4 was low.²⁶ Of those analysed, 7 submitted portfolios (5 passed), 7 submitted written advocacy (4 passed).

Whilst most candidates understood the requirements, and clearly put in the necessary effort to deliver high quality portfolios, where failure occurred it was primarily due to misapprehension as to the detail required or nature of the

²⁶ In addition, 2 candidates' submissions were too late for their results to be included in the data analysis.

document sought. One was a document too short by far to allow of the sort of analysis to which we needed to subject it. The other misjudged the level of detailed explanation necessary to display skills at the top level. The advocate chose a serious offence, but the issues which arose did not allow the advocate to show the degree of analysis necessary to demonstrate skills at Level 4.

5.4.2 A comparison of the assessment mechanisms - Level 4

There was no significant correlation between the two forms of assessment.²⁷ So 4 out of the 7 candidates submitting to Level 4 failed one or other assessment. Only one candidate failed both assessments.

Table 17: Level 4 Results

Portfolio	Written Advocacy
Pass	Fail
Pass	Pass
Fail	Pass
Pass	Fail
Pass	Pass
Pass	Pass
Fail	Fail

²⁷ Kendall's tau-b .091, p= .811

6 Judicial Evaluation

6.1 Background to Judicial Involvement

Judges are those whose work is most closely affected by the standards of the advocates before them. They are also ideally placed to judge the visible aspects of performance in advocacy. It is the judge who is affected on a daily basis by any shortcomings in standards as a whole, and whose ability to ensure that the criminal justice system operates fairly is either strengthened by the skills of the good advocate or eroded by the defects of the poor. Judicial evaluation also has the specific advantage of being based on actual performance, rather than the more artificial assessments of competence posed in portfolios, written assessments and live simulations.

Judicial evaluation was included within the pilot to consider the feasibility of using it as an assessment mechanism (to what extent it could be seen to be fair and transparent and to what extent judges would be able to provide sufficient evaluations to suggest it was a practical option for accreditation). A second reason for including judicial evaluation was to enable the results of JE to be compared with other assessment mechanisms and thereby increase confidence in those other mechanisms.

It is clear from our conversations with them that some judges would welcome the opportunity to be the assessors of advocates before them and others would not. The former welcome it because they do not see how anyone else is in a position to do it, and because of their closeness to the current problems with the standards of advocacy they wish to take responsibility for it. The latter group would not wish to do it because they see the problems relating to transparency and comparability as between the standards used by different judges, as well as the administrative burden which the task would bring. Concerns have also been expressed about confidentiality and the impact of negative reviews on clients seeking to appeal and in particular, whether such assessments might be required to be disclosed. Similarly, many advocates would like to be evaluated by judges on their performance in court, but others have expressed concerns about bias amongst some judges against specific advocates or specific classes of advocate (solicitor advocates, female advocates, and BME advocates in particular).

The use of judicial evaluation poses practical difficulties, and brings with it the risk of unfairness. Trials which are not effective, cases poorly prepared by others and thrust on an advocate late in the day, difficult clients whose changing instructions can only be half guessed at by those not privy to them, can prejudice a judicial evaluation without the judge being aware of the difficulties. To overcome the risk one would need to have candidates assessed by judges on several occasions to be confident that the evaluations were a fair test of competence. It would be possible to monitor judicial evaluations to ensure a degree of consistency between

judges,²⁸ and assessing candidates over multiple performances would mitigate the impact of external factors peculiar to individual cases. The practical issues in endeavouring to use judicial evaluation as part of an accreditation scheme are thereby underlined: sufficient judges would need to provide sufficient numbers of assessments on a range of candidates to enable evaluation to form part of an accreditation mechanism.

As noted above, Judicial Evaluation was sought at Levels 2, 3 and 4. The difficulties of identifying non-lay judiciary able to assess sufficient candidates in the magistrates' courts prevented this approach being adopted for Level 1.

The evaluation instrument at Annex K is a questionnaire designed for use after a trial. Advocates indicated courts in which they regularly appeared, or expected to appear over the pilot period. Those courts were contacted – through their listing office - with a list of the names of candidates who might appear in their court. The list also gave the candidate's UIN, their title, age and professional grouping to assist judges in identifying participants who shared names with other advocates. Those lists were regularly updated. The task of telephoning these court offices to check that the lists had been received and passed on was undertaken by the LSC QAA team. We hoped for one evaluation of each candidate at Levels 2 and 3, with 2 for each at Level 4.

The questionnaire was to be used sufficiently contemporaneously to allow of accurate reflection of performance. It was also developed to be used in the absence of detailed training or briefing, being sufficiently concise to encourage maximum levels of participation from the judiciary. The form was circulated in hard copy and electronically to all the trial centres. The main centres had it from February, the others from April.

6.2 Engagement

The number of candidates who could have been judicially evaluated under the pilot was 148. Of these we received judicial evaluation for 22. Although we welcomed multiple judicial evaluations on candidates (to give a degree of testing of its consistency and reliability) we only received more than one evaluation for three candidates (who each received 2 evaluations).

It was possible that some of the remaining 126 candidates had not conducted trials during the relevant period. We sought data from candidates about the number of trials they had conducted in the Crown Court within the relevant period. Whilst responses were limited, data collected from candidates in respect of whom we had information shows that the 44 candidates who supplied information, conducted 137 trials during the relevant period. We received only 9 pieces of judicial evaluation from those trials. Whilst not all of those would have been assessable (some may not have been heard by full time judges), it demonstrates the difficulties which

²⁸ Such monitoring might not be able to ascertain whether any general biases against particular groups existed.

could attend asking judges to provide evaluations as part of an operational scheme.

It does appear that the more senior the advocate the greater the chance that their need for grading would appear on the judges' radar – and this is recognised in our recommendations. We know that support by judges of the pilot itself has been mixed – some supportive and others antagonistic. Whatever the reasons, we lack significant amounts of data. In an operational scheme where there was plenty of information for grading coming from some parts of the country, and nothing from others, there might be assessment deserts with practitioners in those locations prejudiced.

Setting aside any problems affecting the judges' ability to respond, the information from candidates on numbers of trials show that some have conducted as many as 8 trials in the relevant period whilst others have conducted none, or are still in the one they have been in since January. If an operational scheme were to be reliant on judicial evaluation as the main source of information for grading a candidate this paucity of evidence would mean candidates might have to wait some time before the evidence for assessment was available (and longer still if sufficient numbers of such evaluations were to be built up to ensure their reliability and fairness) and, the candidate might have to be assessed on types of hearing other than trials, which would not enable assessment in all the competences.

6.3 Judicial evaluation of all candidates

For the 22 candidates who were judicially evaluated there were 25 evaluations (3 candidates were assessed by two separate judges). Only 10 of the 22 candidates had submitted to the other assessment mechanisms, hampering our ability to compare the results of judicial evaluation with the other assessment mechanisms.

As noted above, we have only three candidates for whom we can compare the assessments of judges to begin to understand whether different judges mark the same candidates in broadly the same way. This is not sufficient to enable testing of the reliability of judicial evaluation as part of an accreditation scheme. We therefore include the comparisons between the three candidates for interest only. For one of them the judges were in agreement about the candidate. Another candidate one judge rated one point higher on the scales (the difference between a 4 and a 5). For the third the variation was between one and two points on the scales (i.e. assessment on the various criteria range between a 3 and a 5).

Many of the criteria were recorded on 5 point scales 1 = poor, 3 = adequate, 5 = excellent. The following table summarises the average scores on questions with these variables. The left hand column lists the questions posed on the Judicial Evaluation form and indicates their scoring parameters (1 to 5 or Yes/No). N means number (of relevant candidates) and relates to the column to its left.

The 'Both' column indicates the average scores of those candidates both assessed and evaluated and the 'JE only' column shows those who were only judicially evaluated. There are not significant differences between the groups.²⁹

The questions answerable Yes or No were those framed to address more definite concerns about competence. Where answered, the response was generally in the candidate's favour. Two candidates attracted criticism/rebuke from the court but were said to have dealt with it appropriately and courteously. One candidate did not retain the respect of their opponent(s) and the same candidate damaged their case by injudicious questioning. All such unfavourable responses were for evaluation only candidates.

Table 18: Mean scores Judicial Evaluation (assessed and evaluated only)

Question	Mean score (Both)	N	Mean score (JE only)	N
(If familiar with the candidate's abilities as an advocate), before this case , how would you have assessed the candidate's abilities generally? (1-5)	3.7	7	3.9	7
Did the candidate demonstrate a thorough knowledge of the case? (1-5)	4.3	9	4.5	12
Did the candidate have a clear and consistent case strategy? (1-5)	4.1	9	4.2	12
Did the candidate adopt sound tactics consistent with the client's case? (1-5)	3.9	9	4.3	12
Did the candidate adapt to unpredictable developments? (1-5)	4.1	9	4.2	12
Did the candidate remain focused, responsive and effective throughout? (1-5)	4.2	9	4.0	12
Did the candidate (<i>If defending</i>) at all times fearlessly defend the client's interests? (1-5)	4.5	2	4.5	6
Did the candidate (<i>If prosecuting</i>) at all times behave as a fair minded prosecutor? (1-5)	4.3	7	4.5	6
Did the candidate command the confidence of the Court? (1-5)	4.0	9	4.2	12
Did the candidate demonstrate a thorough knowledge of procedure? (1-5)	4.3	8	4.3	12
Did the candidate conduct legal argument accurately and persuasively? (1-5)	4.0	7	4.2	11
Did the candidate show a sound technique in examination in chief? (1-5)	4.1	9	4.4	10
Did the candidate demonstrate a proper understanding of the limits of questioning? (1-5)	4.0	9	4.3	11
Did the candidate adapt the style of questioning to suit circumstances and individuals? (1-5)	4.0	9	4.1	11
Did the candidate cross-examine effectively? (1-5)	3.7	9	3.9	11
Did the candidate deal effectively with expert witnesses? (1-5)	4.0	5	4.7	3
Did the candidate adopt a clear and logical structure for any address to the jury? (1-5)	4.1	9	4.3	12
Did the candidate avoid distorting the facts or evidence? (1-5)	4.3	9	4.5	12

²⁹ Anova, $p > .05$ for each variable.

6.4 Comparing Judicial Evaluation with other assessment mechanisms

In this section we compare the results from judicial evaluations with the other assessments. Given the numbers of assessments where the judges also evaluated a candidate this analysis is limited in nature.

6.4.1 Judicial Evaluation of Level 2 candidates

Of the 10 assessed by assessors and evaluated by judges, 2 were assessed at Level 2. Both failed two assessments. One had failed both cross examination and examination in chief. The other had failed examination in chief and the “must-know” Multiple Choice questions for Level 2.

Both candidates were assessed by the judges as a 3 (adequate) on their abilities generally, with the judges stating the assessment had reaffirmed their view for both candidates. On the remaining specific criteria one candidate scored mainly 4s (and occasionally a 3). The other mainly scored 5s (excellent). These scores do not however appear to tie in with the overall impressions of their abilities at the end of that trial as stated by the judge at the beginning of the form (i.e. a score of 3). This is a small example of the potential difficulties attendant upon requiring judges to evaluate advocates without training in the use of the forms.

6.4.2 Judicial Evaluation of candidates assessed at Level 3

Two of those taking Level 3 assessments were assessed by judges: One had failed the written advocacy assessment and passed on cross examination and portfolio and was given 3s and 4s with some 2s under the witness criteria by the judge. The other had passed all three assessments and was given all 4s by the judge in question. Both were adjudged as adequate overall by the judges, despite the former appearing to perform inadequately on some criteria.

6.4.3 Judicial Evaluation of candidates assessed at Level 4

3 of the Level 4 candidates had been assessed, as requested, by 2 judges. However none of these 3 had themselves submitted any written assessment. Interestingly the other 3 of the 6 Level 4 candidates assessed by judges as well as by the other mechanisms failed on either the written advocacy submission or the portfolio. One failed on both. Those candidates who were also assessed on their own written evidence, mainly got 4s and 5s in Judicial Evaluation, and so were assessed as performing well by the judges, with one getting 5s and one without much judicial evaluation data, due to the nature of the candidate’s practice. The judge used the last trial in which the advocate appeared, which took place almost a year prior to the evaluation. The judge wrote positive comments about the candidate’s performance, but indicated that he was not in a position to grade the performance using the detailed questions.

We compared the only 2 occasions where the assessment team were reviewing the same case as a judge. The evaluation mechanisms came to the same result. A strong performance viewed from the bench was related by the advocate in a way which attracted marks of a similar level from the assessors as from the judge.

6.5 Judicial Evaluation in an operational scheme

The advantages of Judicial Evaluation are that advocates are being judged by actual performance on real cases, and by those charged with exercising some of the most crucial judgments made in the criminal justice system. It does not require advocates to undergo simulation and, if practical problems could be overcome, might therefore be less of a burden on practitioners. Whilst Judicial Evaluation might reduce the overall costs of assessment it would also shift some costs from the profession and onto the court system.

It is also a method which commands the support of a significant proportion of advocates and the judiciary. Such confidence predates of course any testing of the practicality and reliability of the approach. We have not been able to assess reliability in any meaningful sense, but our data does give significant cause for concern about the practicalities of instituting a scheme of judicial evaluation which can contribute meaningfully to QAA. This is likely to include the need for training in the grading of aspects of advocacy performance; the independent recording of the performance [transcripts] and the provision for appeals consistent with other mechanisms; and the need to ensure judicial evaluation is based on data from a number of the advocates' cases to ensure a full range of competences are tested, that assessments were not prejudiced by idiosyncratic cases/clients and to enable some monitoring of consistency of approach across the judiciary. Additionally: it would ordinarily be important for advocates to receive feedback on their performance to enable future development and improvement.

7 Comparison of different professional groupings

7.1 Assessment results by professional grouping

This section of the report considers whether there are variations in assessment scores related to occupational grouping, which is relevant to any consideration of passporting. Differences in quality might be expected to occur given the different professional backgrounds of the candidates in terms of the extent to which they were trained in, and have subsequently specialised in, advocacy at the level being assessed. The data also makes interesting reading given historical controversy over rights of audience. The following tables summarise the results under each assessment by professional grouping. We have done this for Levels 1 and 2 where there are a reasonable number of assessments. Whilst the tables may be read as suggesting some differences, none of these differences appears to reach statistical levels of significance nor did they near significance.³⁰

Table 19: Percentage failure rate of Live Day assessments by professional grouping (Level 1)

	Professional Grouping					
	Barrister	CPS	FILEX	Pupil	Solicitor	Total
Interview	14.3%	0.0%	33.3%	0.0%	10.0%	11.4%
Submission	14.3%	0.0%	33.3%	33.3%	10.0%	14.3%
Cross examination	0.0%	0.0%	0.0%	33.3%	5.0%	5.7%
MCT	42.9%	0.0%	33.3%	66.7%	10.0%	22.9%
Must-know MCTs	28.6%	50.0%	100.0%	66.7%	35.0%	42.9%
N	7	2	3	3	20	35

Table 20: Percentage failure rate of Live Day assessments by professional grouping (Level 2)

	Barrister	CPS	Solicitor	Total
MCT	25.0%	40.0%	57.7%	48.7%
Cross examination	37.5%	80.0%	46.2%	48.7%
Examination in Chief	37.5%	80.0%	38.5%	43.6%
Must-know MCTs	25.0%	60.0%	50.0%	46.2%
N	8	5	26	39

The differences between the different occupational grouping were not statistically significant and there does not thus appear to be a justification for permitting passporting on the basis of any particular occupational grouping, given the level of failures across each of the groups.

In considering these overall fail rates by profession, it may also be helpful to remember the 'over achievements' noted at 5.1 above. Simply put, Level 1 in the framework is defined not by the standard of the advocacy, but by the venue. Of

³⁰ Chi-square tests p for all comparisons all > .1

course it may be said that since many of the cases disposed of in the magistrates' court are less serious, less skilful advocacy may be accepted or even expected.

The advantage to those whose formative practice is in the Crown Court is that it affords the direct possibility, (which a magistrates' court practice does not), to progress to more complex cases. Practice is a necessary part of progression, and any operational scheme needs to recognise this and allow for limited opportunities for advocates to test themselves at the next level in order to progress, if they have the ability.

The evidence calls into question the value of the Higher Rights Qualification as a passporting mechanism for Level 2. Candidates holding it did not appear more likely than other solicitors to cross examine at the standard set for Level 2. This is a matter on which a future scheme might continue to gather data to test the findings of this pilot on a wider group of non-volunteer candidates. Interestingly, this set of solicitors had a lower average number of years' experience (14.6) than those at Level 1 (17.4), whilst the barristers leapt from 1.9 years at Level 1 to 14.6 years at Level 2. It is to be expected that the solicitor contingent at Level 2 has had much less of its practice in the Crown Court than their barrister peers.

7.2 CPS External Prosecutor Grading

This table compares the level at which our candidates were assessed with their CPS external prosecutor grading.

Table 21 Assessment Level numbers by CPS (EP) grading

		CPS External Grading					Total
		No grading	1	2	3	4	0
QAA Level	1	32	2	1	0	0	35
	2	34	1	5	2	0	42
	3	5	0	0	8	1	14
	4	3	0	0	2	2	7
Total		74	3	6	12	3	98

Of 98 candidates assessed, 74 had no CPS external prosecutor grading. Of the 24 with CPS grading, 1 was a solicitor, the rest barristers. 3 had grade 1, 6 grade 2, 12 grade 3 and 3 had grade 4. Two of the candidates assessed at Level 1 also had grade 1 CPS status while 1 had grade 2. At Level 2, 5 of the candidates assessed also had grade 2 CPS status, 1 had Level 1 and 2 had Level 3. At Level 3, 8 candidates also had grade 3 CPS status while 1 had Level 4. At Level 4, 2 had grade 3 CPS status and 2 grade 4. Where candidates whom one might expect to hold a CPS grading (barristers at Levels 3 and 4) did not do so, this may be because the grading system is not a national one and in some areas it is not therefore possible for an advocate to acquire this status.

This table compares the CPS grading with candidates' *performances* in cross examination, thus only those assessed by this method - Levels 1, 2 and 3.

Table 22: CPS Cross examination Level achieved compared with CPS (EP) grading

		CPS External Grading					Total
		No Grading	1	2	3	4	0
Level Achieved on Cross examination	0	3	0	0	2	0	5
	1	34	1	3	2	0	40
	2	28	1	3	0	0	32
	3	2	0	0	3	1	6
	4	2	0	0	2	0	4
Total		69	2	6	9	1	87

No candidate with CPS grading failed at Level 1. However, 1 of those Level 1s holding CPS grade 2 achieved only a Level 1 in cross examination. Of those assessed at Level 2, 3 with CPS grade 2 also achieved Level 2 in their cross examination, however 2 with Grade 2 and 2 with grade 3 achieved only Level 1. It may be that the seriousness of the cases with which these candidates usually dealt led them to underestimate the job to be done in this cross examination. – but that is something which could apply to the Level 2 underachievers generally. Those assessed at Level 3 had the highest proportion holding CPS grading. One held Grade 4 and passed at Level 3. The rest held Grade 3, of whom 3 also achieved Level 3 in their cross examination but 2 failed and 2 achieved Level 4.

8 Which assessments for which level? The possible shape of a scheme

8.1 Cross-checking competence assessment

In making recommendations for the shape of an operational scheme we have reverted to the list of competences we deem certainly assessable by only one of the methods we have used. If our proposals discard a relevant assessment mechanism they will either have to introduce an additional necessary element to a retained instrument or acknowledge that our recommendations do not leave us certain that the specified competence(s) are assessed.

A.1.1	Presents and questions only material witnesses	Pf
A.3.3	Takes appropriate advantage of new material	Pf and LA
B.2.2	Provides appropriate disclosure of evidence	Pf (if arises) more applicable to prosecutors
B.2.3	Obtains instructions when appropriate	Pf
C.3.2	Deals appropriately with vulnerable witnesses (only if criteria made it a technical requirement)	Pf
C.3.3	Deals effectively with uncooperative witnesses	LA (if so designed)
D.1.1	Drafts clear skeleton arguments	WA
D.1.2	Speaks clearly and audibly	LA
D.1.3	Maintains pace <i>throughout the course of the trial</i>	JE (but without italicised part, assessed also in LA)
D.2.1	Questions to witnesses are clear and understandable	LA
D.3.3	Assists the court with the proper administration of justice	MCT
E.1.2	Ensures team members are allocated tasks consistent with their level of competence	Pf

8.2 Place of judicial evaluation

The discussion at 6.5 above illustrates the practical difficulties which give rise to continued doubts about the desirability and feasibility of any scheme of Judicial Evaluation in the absence of greater numbers of such evaluations and the testing of them for reliability. **We cannot therefore recommend it at this stage as an assessment mechanism for the majority of advocates.**

If Judicial Evaluation is to be developed as part of a QAA process **we believe the practical difficulties should be addressed on a cohort of candidates where the numbers are not overwhelming in practical terms and where judicial commitment to the exercise is likely to be at its highest. This suggests concentrating Level 4 assessments, where there is the strongest prospect, initially, of developing Judicial Evaluation into a valuable tool.**

At Level 4 the pilot attempted to obtain from all candidates two types of Judicial Evaluation. One was to be provided, as for candidates at other levels, without the advocate seeking it or necessarily knowing that it had been produced (to avoid

behaviour modification). The other was to be sought by the Level 4 candidate from a judge who had observed their performance in a recent trial. Candidates were asked to provide any such judge with a copy of the form as an indicator of the areas which their evaluation should cover, but to invite the judge to produce an evaluation of a more descriptive nature, explaining where the advocate's strengths and weaknesses lay. None took that opportunity but if this were a requirement of a scheme it may be that they would embrace it, particularly if relieved of the burden of grading the vast majority of advocates.

8.3 Diets of Assessment

A key advantage of simulation over real life assessment is that the simulations ensure that all candidates work with the same materials and are tested on the same facts and issues. Thus an assessment of one candidate is broadly comparable with the assessment of another candidate. The simulation also allows a wider range of issues and tests to be built in than would ordinarily be found in a real case. The simulated advocacy exercise can test the performance skills of the advocate and their ability to identify and use evidential and legal points. With suitably designed assessments, it provides the opportunity to feed the candidate late information to test the speed and accuracy of their response. It also tests their ability to adapt 'on their feet' to a response not in line with expectations, or a witness's particular characteristics. Thus, in important senses, the 'unreality' of simulation can be an advantage: it can be designed to include particular characteristics essential for the testing of competence at the relevant level. It could take several real cases to afford an advocate such an opportunity.

The cross examination assessment assesses a wide range of competences; in a context as similar to genuine advocacy as possible; on a basis which is fair and replicable (candidates sit the same standard of test); and it is not prone to problems that portfolios and other assessment tools are prone to (that they test a candidate's presentational skills as much as they test their actual skills). For these reasons the research team and the assessors had most confidence in this method of assessment. Our own exit feedback from candidates, in common with that in the LSC feedback report, supported this view. Whilst opportunities to compare judicial evaluation with other assessments were limited in number, judicial assessment of a candidate's overall performance was most often the same mark as that which they had given for the effectiveness of their cross examination, i.e. judges appear to see cross examination as the key skill.

A further advantage of a simulated cross examination is that it can be used if required as a gatekeeper to assisting the progress of candidates up the levels. An advocate aspiring to a Level could first (without risk to a real client charged with a more serious offence) attempt to obtain a pass at the relevant level in such an exercise. A fail would have no impact on clients facing charges or victims having suffered crimes at that next level. A relevant pass could however give the right (for a limited period or for a limited number of cases – or a combination of both) to

accept instructions to conduct a trial of an offence categorised at that next level. An MCT could be used in conjunction with this in the same way, as a gatekeeper.

Whilst portfolios can be managed by candidates to enable them to present themselves in their best light (problems more particularly rehearsed at **2.3.2** above), the team of assessors felt that they provided good insight into the level of experience a practitioner has achieved and their ability to reflect on that experience. Interestingly, and to the surprise of assessors, candidates who undertook these were appreciative of a rare opportunity to reflect upon and relate their own practice. It also affords the opportunity of requiring different levels of skill to be displayed consistent with the case in which the advocate is appearing and the level to which they are being assessed. When an advocate is assessed at the higher levels, a reduction in the number of cases required for the portfolio reflects the added complexities of any case with which they would be dealing and the extent to which they would be expected to explore their analysis of that case and their approach to conducting the trial.

8.3.1 Multiple Choice Tests in operation (Levels 1 and 2)

From our analysis a Multiple Choice Test is an assessment without a proxy and is a valuable additional tool which has the advantage of readily enabling assessment of new procedures or interesting points of law or evidence. A disadvantage of a Multiple Choice Test is the regard in which it is held as a tool by some members of the professions. For many of them it is linked to testing of a level of knowledge and understanding consistent with being a student and does not appear to have the sophistication which they hope characterises their current level of competence. They may fail to recognise the particular utility (breadth of assessment, consistency, economy of development and administration) of the tool, and the fact that it can be drafted in a way which tests some particularly difficult points.

The main operational difficulty of a Multiple Choice Test is not in sitting it initially, when a candidate is also taking a cross examination test, but in arranging a reassessment which might occasion either a costly attendance or arrangements for secure remote assessment (by use of the internet) which could entail significant development. An alternative solution might be to engage the professions through their local presence to administer any such test. However the need to test D3.3 could be handled with a lighter touch by the portfolio requirement.

One aim of the research and evaluation conducted was to reduce the number of assessments to the minimum which could properly assess the competences. Interviews (or conferences) and submissions were shown by the data to give results which largely correlate. There are aspects of the competences – concerned with relating to the client - which can only be assessed live through interview, but may be judged in part from a portfolio.

There is one competence, **A.1.3** (*Making only relevant submissions*), which we consider capable of assessment in the recommended regime of assessment for Levels 2, 3 and 4 only through Written Advocacy. Written Advocacy does not appear as an assessment instrument in the recommended diet at Level 1, where

the research and feedback indicates that it is not appropriate. We thus consider that, if this competence is to be assessed at Level 1, then the Submission advocacy assessment be retained at that level only. The alternative would be to dispense with that competence at Level 1 and thereby the need for the Submission advocacy. The interview can, we suggest, reasonably be dispensed with because of the correlation between it and the submission in the data, and because of the support for its specialist areas of competence which the portfolio can give.

8.3.2 Assessment Recommendations

As a result of this pilot, we believe that an assessment landscape which balances the need to assess a range of competencies with the proportionate testing of those competences can lead to a simpler assessment regime than that trialled in the pilot:

Level 1

A portfolio to include 3 trials at Level 1; at least 2 to be of either way offences.

A submission based piece of advocacy

A cross examination.

and

Specifically if **D 3.3** (*Assists the court with the proper administration of justice*) above remains a competence, then **either** a Multiple Choice Test with a relevant question or a requirement in a portfolio case to show that competence would be needed. However, since **B 2.1** (*Complies with appropriate Procedural Rules and judicial directions*) remains at this Level an untested competence, a preference for the Multiple Choice test option would enable capture of **both** competences.

The suggested number of trials for the portfolio at this Level reflects the fact that the trials will be shorter, thus easier to relate, and that their reduced complexity suggests this number is more likely to ensure the competences are satisfied.

No written advocacy is recommended at this level due to the fact that our cohort was rarely called upon to do this. They could not provide such evidence in any realistic way. The need for this could be kept under review should practice change in the magistrates' court.

Level 2

A portfolio to include 2 trials at Level 2 and a piece of Written Advocacy.

A cross examination

and depending on the view taken about capture of competences D 3.3 and B 2.1, as set out in relation to Level 1, a Multiple Choice Test.

Level 3

A portfolio to include a single Level 3 trial and a piece of Written Advocacy.

A cross examination.

To ensure that competence **D.3.3** (*Assists the court with the proper administration of justice*) is captured at this Level it would be necessary to include a requirement that the case described in the portfolio included an instance where the advocate did so.

Level 4

A portfolio to include a single Level 4 trial and a piece of Written Advocacy.

A cross examination.

At least 3 pieces of Judicial Evaluation:

Two of these on Judicial Evaluation forms, one covering the same case as in the portfolio (the advocate would have to alert the Assessment Organisation so that the court could request that the judge do the evaluation); a second from another case, and a third a narrative document detailing the advocate's strengths and weaknesses and provided at the request of the advocate by a judge (perhaps of High Court level) who has been the judge in a trial lasting at least 3 weeks in which that advocate appeared. The reason for recommending three pieces of judicial evaluation is to reduce the potential impact of any inconsistent or unfair evaluation.

If competence **D.1.3** (*Maintains pace throughout the course of the trial*) were modified by the deletion of the words in italics above, it could be assessed in the above regime through cross examination. If not it would remain unassessed.

If such a scheme were to become operational it would leave only 3 of the competences not already recommended for deletion with any uncertainty over their being assessed. These are:

C.3.1 Observes restrictions and judicial rulings on questioning

D.3.1 Observes duty to the court and duty to act with independence

D.3.2 Advises the court of adverse authorities and, where they arise, procedural irregularities

Each could be satisfied by a requirement (in the guidance to candidates on portfolios) for them to choose cases which showed particular challenges on these points, but to do so would restrict the cases from which an advocate could choose to those which showed particular problems relating to these competences. If not they too could be dispensed with.

8.4 Levels and movement

Throughout the report we have commented frequently on Levels. Proper grading of advocates according to recognised Levels is at the heart of the QAA scheme. Achieving the ability to distinguish cases readily by Level will be essential – but this will never be an exact science. This lack of strict boundaries (save at Level 1) can be used to advantage in an operational QAA scheme. It will mean that even without formal recognition of a “ticket” at a higher Level, advocates will be able to take on cases at the upper margins of the Level at which they currently have the right to practice. This should enhance their skills.

This lack of a precisely drawn boundary can also assist the scheme in another way. The advocate aspiring to the next level could take on responsibility for a limited number of cases at that higher level for a limited period without having to make such a big leap. It could be a recognition built into the scheme which acknowledges that an advocate does not arrive fully fledged working at a level.

Our recommendation is that there should be a way of testing some competences before any entry at all to a Level (a gatekeeper assessment) and that thereafter, until they have achieved passes in the full diet of assessments an advocate should be subject to a probationary period of 12 months within which to acquire the remaining pass in the non gatekeeper assessment – e.g. the portfolio.

Failure to pass the full set of assessments within the probationary period would give rise to the advocate having to recommence the process. In this way those people actively seeking to improve their knowledge and skills at the next level will have an opportunity of so doing. If anyone does not take that opportunity, there will be a restricted period during which they will have access to those higher level cases.

There is a particular problem with this approach which stems from the current drivers to ensure that an advocate takes responsibility for a case from an early stage and continuously. In the past, one way that an advocate became familiar with a higher court, and improved their own ability to deal well with matters in that forum, was by taking responsibility for earlier hearings of matters which they would not conduct at trial. This will still happen in earlier stages of the most serious cases, but for many others it will not.

The responsibility is therefore very much on the advocate to take on cases of incremental difficulty within the Level. This is more difficult in an era where senior colleagues or clerks are keen to see that the advocate is as fully and profitably engaged as possible. However the codes of each profession are clear on the need for an advocate only to deal with cases within their competence. The onus remains as ever with the individual to resist stretching themselves too far.

8.5 Practical problems and suggested solutions

8.5.1 Defining criminal proceedings and defining advocacy for QAA purposes

Whilst the QAA pilot has focused on advocacy in criminal proceedings, there is currently no definition of criminal proceedings nor of advocacy. This has not given rise to any practical problems in the pilot but is something which may be more likely to give rise to problems as the scheme develops and is applied to decisions on the funding of real cases and accreditation of candidates.

8.5.2 Defining Levels

Candidates appeared to struggle to define the level at which they should be assessed. There was only one assessment instrument which allowed us to gauge whether the candidates had “correctly” categorised the level of their real practice - and that was the portfolio. Their level of practice was dictated not by anything readily identifiable about them as individuals (e.g. years’ practice, qualifications held), but by the cases they dealt with. This meant that an advocate’s ability to select their level depended on their ability to interpret the Levels Document (Annex B). Even if advocates were clear about the level at which they wished to be assessed, they were sometimes unsure of the level of their current cases. If a volunteer cohort found drawing those distinctions difficult, it can only be supposed that this problem would be common to others.

The uncertainty outlined above was not confined to lower level cases. In particular, the assessors marking (and the advocates undertaking) the high level cases felt that there was no easy divide to be made between Levels 3 and 4. In fact 2 serious cases of a sexual nature written up for the portfolio submission, one at Level 3 and one at Level 4, had nothing to differentiate them in Level – save the perception of their **own** level of the candidate submitting them. The fact that sexual offences are to be found at all three Crown Court Levels without clear guidance as to where the divide lies will give to this category of offence a particular uncertainty. This is related not by way of criticism of the candidates but to point out the inherent problem of the levels as currently framed.

We recommend that particular consideration is given to defining sexual offences more specifically if the Levels as they currently stand are to work efficiently in an operational scheme.

Producing a four Level table, capable of easy digestion and reference, for the multiplicity of offences and manifestations of those offences had been a challenging task for the Levels Group, and for the pilot we used the levels drawn up by the Levels Group. However, when researching prior to starting the pilot, the research team had devised an alternative way of identifying levels of case. We did so because we feared that there would be difficulty for advocates, those instructing them and their paymasters, in identifying correctly the level of a particular case, or some categories of offence. Having used the levels provided, it appears that at least for some offences, and for this cohort, the problems we anticipated

materialised. Given the relatively few cases dealt with in the pilot, the problem could only be expected to increase in an operational scheme.

We recommend that the Levels Document is revisited with a view to clarifying further the offences and their gravity which are to be found at each Level.

It may be that the structure we devised merits further consideration and in case it does we attach it as Annex L. But even if this is not considered it will be necessary in any operational scheme to accept that before such a scheme is embedded – and even to a certain extent afterwards – there will be practitioners straddling the levels. Indeed this fact could be used in a structured way as suggested at 9.4 to provide the sort of flexibility in the scheme which will ensure that the courts do not suffer for lack of an advocate at a particular level.

8.5.3 Ensuring breadth of practice

It is acknowledged that one of the central ways in which an advocate will improve in competence is by conducting more and increasingly difficult cases. The suggestions about diets of assessment made at 8.3 envisage a portfolio limited only to trials at the relevant Level, and only to 2 or at most 3 of those. A portfolio of this sort was used in the pilot because we were testing competences rather than the breadth and depth of a practitioner's experience at a particular level. However it would be possible in any operational scheme for there to be a requirement for an advocate to log the numbers and types of hearing they have conducted in a stated period prior to submission of the portfolio. If this were thought of benefit (and we suggest this has particular relevance at Levels 1 and 2) there would have to be technical requirements about its content. But that would be akin to the current requirements for Duty Solicitor accreditation.

8.5.4 Probationary status

If the recommendation of gradual progression into a Level were accepted, **there would have to be Registers of Probationers at each level to ensure proper monitoring of the QAA scheme and the proper attribution of costs to cases.** Such registers already exist for police station and Duty solicitor work and the LSC currently has responsibility for them through First Assist. A body responsible for managing such a QAA register would similarly have to be set up, or an existing body take it on.

8.5.5 Reaccreditation

Any consultation on QAA should consider the timing and process for reaccreditation (retaining accreditation at the current Level). There is a need to ensure that reaccreditation processes pick up changes in competence over time with the need to ensure candidates are not affected by a disproportionate assessment regime. A number of candidates have also recognised the need for reaccreditation in their feedback comments.

We recommend that a new QAA scheme should not require any reaccreditation before 5 years from the time of an advocate's first accreditation, though there may be some support for having a period of 7 years. In the shorter time frame many advocates in the earlier part of their career will in any event have sought accreditation at the next level. If they have not it would appear prudent to reassess, but to give credit for their earlier display of competence by a reduced diet of assessment.

The portfolio would appear to be the cheaper assessment to offer and may for this reason be the more popular proposition among some practitioners. It would be readily prepared by an advocate already working at the Level from which the necessary case(s) would derive. It does not however have the rigour of a live assessment, which we believe is the best assessment, has the strongest support among the pilot cohort, and does not require the same length of time to prepare. The actual expense of taking a live assessment is greater than a portfolio if outlay alone is considered, but the length of time necessary to prepare a competent portfolio makes it at least as expensive in its hidden cost. We suggest therefore that a cross examination exercise is likely to afford the degree of reassurance about the standard of an advocate's practice which is necessary for reaccreditation.

If a portfolio were to be the reaccreditation method adopted, there could be an added requirement that an advocate failing a portfolio based reaccreditation be reassessed by a cross examination assessment (an incentive to submit a properly prepared portfolio).

9 Passporting, Exemption and Assessment in an operational scheme

The following terms appear in this part of the report:

Passported means automatic admission to the level without any need for assessment. When a candidate is described as passported it is by virtue of previously acquired qualifications or status.

Exempted denotes candidates being exempted from one or more assessments at the relevant level.

The team was specifically asked to consider whether and where it would be possible to make recommendations for a reduction in the number or need for assessment of any group of candidates based upon their previously acquired status or qualifications. The relatively low number of candidates assessed makes it hard to recommend any passporting in respect of already held qualifications.

The lack of significant variation between the performances of different professional groupings adds to this difficulty. However we do make recommendations where there is other information which we believe leads to a defensible approach to passporting or exemption, with the proviso that the scheme be subject to review.

We recommend that appropriate exemptions (which could include exemption from all relevant assessments) be given to participants based upon successful performance in the pilot. This is in addition to any recommendations which we make in respect of the particular professional groupings discussed below.

Specific recommendations at each Level are set out below.

9.1 Level 1

Pupils

There were only 3 pupils - their number and variations in performance mean that no recommendations for passporting can be made. Two failed only the Multiple Choice Test, one failed the Multiple Choice Test and the advocacy. Their performance levels in the cross examination ranged from Fail to Level 2. All had achieved a Very Competent for their advocacy on the BVC. However, once a pupil has the certificate of successful completion of the first six months of pupillage, that person holds the same rights of audience (subject to the usual restrictions imposed by the professional code) as any barrister. We therefore recommend that this group be treated in the same way as barristers at Level 1 (see below).

FILEX

There were only 3 Fellows of the Institute of Legal Executives. Their number means that no recommendations for passporting can be made based upon the

data. All passed the cross examination and 2 passed all assessments. None performed above their relevant cross examination Level.

It should be noted as regards Level 1 advocacy accreditation, that this group of advocates has had to pass up to 8 different advocacy assessments to achieve their status as Fellows with rights of audience in criminal proceedings in the magistrates' court.

Solicitors

Most solicitor candidates passed the Level 1 assessments. Two had fails in more than one assessment, 14 passed all elements they took. Only 1 candidate (2 if the Multiple Choice test is retained) would have failed the diet of assessments we recommend at section 8.3.

There is no evidence to show that Duty Solicitors performed any better at Level 1 than solicitors who were not Duty Solicitors. However the Level 1 cohort of 20 solicitors consists of only 4 who are not current Duty Solicitors – and one of them was previously a Clerk to the Magistrates and Duty Solicitor. The vast majority of that cohort passed the instruments which we recommend as part of an operational scheme. One failed the cross examination and 2 the Multiple Choice Test.

It could be argued that this cohort is likely, where the majority are Duty Solicitors, to include those with significant exposure on their feet as advocates in the magistrates' court. They have already been scrutinised in order to achieve that Duty Solicitor status. The assessment instruments used for Duty Solicitors lack any assessment of witness handling or of knowledge in a Multiple Choice Test. Nonetheless this cohort has shown itself largely competent on those tests. This may be because of the exposure they have gained. It should also be noted that Level 1 candidates were volunteers and there are plausible reasons for speculating that they might appear to be of a higher quality than a random sample of candidates.

A suitable balance should be struck between recognising their prior accreditation and assessing their quality.

We therefore recommend that Duty Solicitors should acquire Level 1 status with only the need to pass the cross examination exercise.

Other solicitors would need to take the full diet of assessment.

Another matter on which we would make a recommendation is in respect of all solicitors for whom criminal practice forms the main part of their work, and has done for a considerable period. It may be that, in recognition of this, any future scheme could trial giving exemptions on all but the cross examination exercise to any solicitor with a criminal practice with a minimum number of years' (say 10) experience in this field.

Barristers

Eight barristers who were not pupils were assessed at this Level (2 of them work for the CPS). Three barristers failed the Multiple Choice Test, but none failed the cross examination. Both CPS candidates passed all assessments. Only 1 barrister had more than one fail. The numbers are insufficient to make wholesale recommendations based solely on this data as to passporting or exemptions. It should however be remembered that this cohort had on average only 1.9 years' criminal practice. **We recommend consideration be given to allowing a barrister exemption from any cross examination and submission test at Level 1.** This would be an appropriate recognition of the training and assessment they have undergone on the BVC. To do this would also sit most consistently with the findings about barristers at Level 2.

If a Multiple Choice Test were part of the diet of assessments there is no basis upon which to suggest any exemption be offered in respect of this assessment.

We therefore recommend that the Multiple Choice Test be used as the gatekeeper assessment for barristers at this Level. Passing the Multiple Choice Test would trigger a 12 month probationary period within which a portfolio must be passed.

9.2 Level 2

Solicitors and barristers who presented for assessment at Level 1 overachieved to this extent – 2 of 7 barristers and 8 of 20 solicitors assessed at Level 1 achieved Level 2 in their cross examination. However, the cohort at Level 2 looked rather different. It contained – for the Live assessments 39 candidates of whom 26 were solicitors, 8 barristers, and 5 held either professional qualification and worked for the CPS. The average years' criminal practice was 14.6 for solicitors (less than for those assessed at Level 1) and 11.7 for barristers (much more than those assessed at Level 1). These solicitors underachieved in greater numbers and to a greater extent than the barristers. In the cross examination 11 of 26 solicitors failed (42.3%) while 2 of 8 (25%) barristers did so. Only 1 CPS candidate cross examined at the correct level.

What the evidence from the Level 2 assessments suggests with some force is that there is no evidence to show that merely having the right to appear and conduct trials in the Crown Court means that an advocate can exercise their skills at the requisite entry Level for that court. Currently barristers automatically have that right and solicitors can acquire it. Feedback evidence shows that even when solicitors have that right they may not have much experience of trials in the Crown Court.

We recommend trialling a full diet of assessments for entry to this level coupled with a reduced diet for those already working at this level, with a minimum requirement of a cross examination assessment for all to be passed within a year of the inception of any operational scheme.

9.3 Level 3

Numbers assessed at this level were too small to be able to make recommendations for passporting. 13 advocates (4 solicitors and 9 barristers) were assessed via a Live Assessment Day, 10 candidates also submitted a portfolio. There was no difference between the failure rates of the professions – 1 solicitor and 2 barristers failing. 3 barristers and 1 solicitor assessed at Level 3 were found to be performing at Level 4 in the cross examination exercise.

We recommend that a new entrant at this Level take a full diet of assessments, and for those whose practice is already at this Level the requirement within 2 years of the inception of an operational scheme to have passed that diet of assessments. The greater length of time takes account of the fact that those practising at this level will have fewer and longer trials.

9.4 Level 4

Numbers assessed at this Level (7) were too small to be able to recommend passporting. Only barristers were assessed at this Level.

Some Level 4 candidates contacted the principal investigator to give feedback on the experience and said they would have welcomed the opportunity of undertaking a live assessment. Our development and testing of the Level 3 cross examination leads us, and our assessors, to believe that this would be possible.

We therefore recommend the addition to the diet of assessment piloted at Level 4 a cross examination exercise of the same degree of complexity as that taken by Level 3 candidates, but assessed to the more exacting standards expected of a Level 4 candidate.

This exercise affords the possibility of achieving a pass at Level 3 or 4 for Level 3 candidates, providing an indication of their readiness to proceed to other Level 4 assessments and facilitating progression through the scheme.

We recommend that a new entrant at this Level take a full diet of assessments, and for those whose practice is already at this level, the requirement within 2 years of the inception of an operational scheme to have passed that diet of assessments. The greater length of time takes account of the fact that those practising at this level will have fewer and longer trials.

9.5 CPS External Prosecutor Grading

The data at section 7.2 gives no reason for passporting those with CPS grading into a particular QAA Level, save by virtue of their performance in the pilot. The fact that the CPS grading system is not standardised is supported by the evidence. Neither is it applicable countrywide. We not to recommend any exemptions based on CPS grading.

10 Setting up an operational scheme

10.1 Establishing a framework for running a QAA scheme

There are many challenges to the successful establishment of an operational scheme. There needs to be acceptance by all professional groupings of the need for QAA. Even a determined judiciary and an adamant paymaster need co-operation for what would be such a radical change to take place. Our feedback shows that among the volunteer cohort the vast majority – for a variety of reasons – think QAA is essential. Even those who feel they are not yet performing up to the standard of the court in which they have rights of audience, have expressed their desire to put themselves through assessments (and any necessary training or experience) to get them to that point. Very few, it appears, are content with the current contoured landscape of standards, deriving as it does from a patchwork of profession specific qualifications, and underpinned by experience gained from instructions obtained sometimes, but not necessarily, on merit.

There appears little confidence that the separate schemes (where they exist) are achieving the proper degree of rigour or comparability in results. Nor have such schemes established common standards across professional groupings. It is to be hoped that a scheme subscribed to by all groupings and their professional bodies would give confidence to all advocates that they could rely upon a fair career progression as an advocate, based upon their skills and experience and not upon which professional grouping they belonged to, or whether or how they had been assessed since qualification. Similarly, understandable concerns about new initiatives can be minimised by the professional groupings working together in establishing assessment processes and standards.

The key to addressing this problem could be a cross professions authority (one candidate called it an Advocacy Board) which would have responsibility for the establishment of a scheme, regulations, validation and periodic monitoring of assessment organisations who would run assessment diets for which they were authorised.

Another difficulty is that those with vested interests in existing schemes may wish to see those separate schemes continue. Appropriate recognition of any qualification shown to be valid for the new scheme could address part of this problem. The gradual replacement of the patchy landscape by a contiguous one could free up resources currently expended by all the professions and the LSC in the administration of those separate schemes. This could add to the funds available to run QAA, for which substantial funding would be necessary for the early work of validation and later on in ensuring comparability of standards as between assessment providers.

10.2 Who is to assess?

10.2.1 General

It is essential that the professions have confidence in a process of grading which has the potential to affect the careers of their members. In order to ensure this, it is envisaged that once an operational scheme has been accepted it will be overseen by the professions – preferably operating via the sort of joint committee mooted above.

It is hoped that this report's description of the processes needed to ensure fairness, transparency, consistency and rigour will assist in establishing an operation scheme. Assessment organisations would be required to demonstrate that they had suitable experience and ability to offer assessment (and where appropriate training) using the standards to which the pilot operated and the methods recommended according to a set of regulations to be devised in the anticipated consultation.

10.2.2 What kind of assessor and assessment organisation?

Comment has been made in some feedback about who should have the responsibility for assessing. The experience of the pilot is that successful and consistent assessment relies on a variety of skills which are not to be found solely in either those now working in education or current practitioners.

The newest of our assessors - two of the Level 3 and 4 practitioner assessors recruited for the pilot - would readily admit that they relied upon the experienced academic assessors for the preparation of assessments and criteria, and the design of assessment days. In addition they required the guidance in using criteria, as well as operating fair assessment days which those who are expert in this provide. Equally, those academics involved in the process – all either barristers or solicitors and previously advocates - who do not still practice, recognise the need to have a constant dialogue with current practitioners about standards, practices and practicalities of the skill in its current exercise.

A successful scheme also needs good administrative support to ensure that assessments are arranged and run within the sort of framework which ensures fairness and allows an advocate to work their practice around the assessments. This means that the ideal is to have assessment carried out using a team of practitioners now in education and current practitioners but backed up by the administrative support of a single institution or organisation experienced in assessment. There are in existence around the country centres where such bodies of experience either exist currently or could be gathered in order to provide a sufficient pool of assessors.

10.3 The logistics of assessment

The report has outlined some of the difficulties in providing appointments for assessments (see Annex L), or windows within which the cohort was able to submit written assessments. Our experience in other compulsory schemes leads

us to believe that these problems would not beset an operational scheme to the same extent because candidates would no longer be volunteers. Nevertheless any scheme will need to be mindful of the pressures of practice.

Assessments must be provided in a rolling and repeating programme with sufficient frequency to enable ready access to opportunities for an advocate to either obtain at first grading at the relevant level and eventually to progress to the next level.

Those assessments need to occur around the country at sufficient centres to make them accessible to all advocates. A balance however has to be struck on the number of centres. A proliferation of centres increases the cost of the assessment and reduces the frequency at which assessments can be conducted. It also increases the difficulties of monitoring and ensuring consistency. The provision of those assessments by too many assessment organisations would make it difficult for the authorising body to properly monitor their systems and standards. This might leave a new scheme facing the same difficulties which beset the separate schemes which currently exist, if that body cannot ensure the same degree of rigour and consistency applies to all assessment organisations.

Our experience in offering nationwide assessment to Duty Solicitors is that advocates are more concerned to have an assessment on a specific date or in a particular time frame for admission to a rota than to have one in proximity to their home or work.

The assessments for the QAA pilot were offered on different days, including Saturdays, in the daytime and even some evenings. Appointments were given to fit in with childcare and court commitments. This sort of flexibility can only be achieved with very committed assessors. Evening work will create extra cost as in some centres it will necessitate an overnight stay. Those we offered were only taken up as late as 5.30pm. As between weekdays and Saturday there was a slightly lower rate of acceptance of appointments for Saturday, but this would not necessarily be the case for an operational scheme.

10.4 Monitoring and ensuring consistency across assessment organisations

The greater the number of assessment organisations validated, the harder it is to monitor comparability. Should such a model as described above be adopted there are 2 main ways in which this could be safeguarded. The first is that there could be monitoring of the assessment organisation during the first year of their delivery of the scheme; the monitoring to be undertaken by such overall board as is set up to administer QAA. A system of monitoring could then be implemented, including requirements for statistical reporting on assessments and candidates, which decreased in frequency as confidence grew in assessment organisations, but increased if problems arose.

The second way to ensure continued fairness would consist in annual joint meetings of representatives of the markers at each level from all authorised

assessment organisations, where sample assessments could be compared and guidance given as to where the appropriate standard should lie.

This is of particular importance given the atypical nature of the candidate sample in the pilot. It would be sensible if a system of QAA is implemented on a compulsory basis for a period of continued testing, refining and monitoring to take place as the scheme beds down. There would also be a need to consider making recommendations on double-marking as the range of assessors increases in any live scheme.

10.5 Feedback

Candidates participating in the pilot do not receive results and thus no feedback has been provided to them. Best practice would indicate that feedback is desirable for all candidates. For reasons of practicality and economy feedback is generally provided in existing schemes only to those candidates who fail an assessment, or to those whose performance or submission gives cause for concern despite the fact that they pass.

We recommend, consistent with current practice, that candidates are provided with feedback on their performance in terms appropriate to each of the assessment mechanisms used.

10.6 Dress

Normal practice in advocacy accreditation processes is to require advocates to present themselves for assessment dressed appropriately formally for court but without being robed. No such requirement was made for the pilot, but

we recommend that any QAA scheme adopt the practice of appropriate formal dress.

11 Costs of an Operational Scheme

We anticipate that the costs of obtaining accreditation will be a significant concern for practitioners, and this is reflected in the Commission's summary of candidate feedback (para. 2.9.1). Therefore, we have endeavoured to estimate the likely cost to candidates of completing accreditation at each of the Levels. The costings draw on both our experience of running the pilot assessments, and our experience of running schemes which include a similar diet of assessments, such as CLAS (Criminal Litigation Accreditation Scheme). It is confined to an assessment of the costs of conducting the assessments, and does not include the costs of implementing the Scheme as a whole.

Any estimate of costs has to make provision for:

- developmental costs (particularly preparation of materials);
- administration costs;
- overheads (such as venue charges);
- assessors' fees (including live assessment days, marking of portfolios/written advocacy and attendance at Test Boards);
- actors' fees and fees for expert witness for cross examination assessments at higher Levels;
- fees for external examiners.

If the QAA scheme is to be offered via Assessment Organisations, then those organisations will wish to secure a financial return, so an uplift to reflect profit margin would also need to be applied. The amount of that uplift will vary depending on the nature of the organisation, the extent to which it is able to make any economies of scale, the administration system already in place and the extent to which it can tie in the development of this scheme with any others it has run.

The amount to be charged will also vary depending on whether or not the organisation has to charge VAT. All costs and likely fees are quoted exclusive of VAT and are conservatively drawn – taking no account of the factors above which might operate to reduce either the suggested costings or the impact of the uplift applied in respect of profit.

It should be remembered that some of the advocates already pay significant amounts for their professional accreditation. If a new scheme either wholly or partially replaces any such accreditation there would be either total or partial saving of those fees.

Additionally it would be possible for the professional regulatory bodies to recognise any accreditation - or reaccreditation – under QAA for the purposes of the requirements of compulsory professional development. This would both ensure that CPD undertaken was entirely relevant to the lawyer's area

of practice and would also have the effect of reducing the real cost to an individual advocate of QAA assessment.

11.1 Candidate Fees – Level 1

We estimate the average costs of providing a Live Assessment Day at Level 1 as:

Venue Hire -	£700
Assessor's Fees -	£700 (based on £600 fee plus £100 travel expenses)
Actor's Fees -	£250 (based on £150 fee plus £100 travel expenses)
Administrator -	£100

Therefore, the costs of running a Live Assessment Day at Level 1 would be in the region of £1,750. We anticipate that a maximum of 8 candidates could be processed in a day by such a team. It would be possible to assess more, but the practicalities of filling such an assessment day, as well as the pressure which a fuller day puts on assessors lead us to use this figure. With such numbers the figure per candidate would therefore be £220.

To that would have to be added the costs of assessing a portfolio and marking the multiple choice test which for this purpose we set at £60 and £10 respectively.

Additionally, the fee charged to candidates would also have to provide for developmental costs, administration costs and the costs involved in maintaining an external examiner system/assessment boards to review results. By way of guidance we would expect each single set of case study papers as well as an MCT to cost (including developmental meetings) in the region of £2,000. Case studies and MCTs have to be used cyclically and revised. They can therefore only be used a limited number of times. We estimate that the approximate cost per candidate in respect of all the above would be £110.

Therefore, the costs per candidate at Level 1, before application of any uplift to reflect profit margin, would be approximately £400. To make the scheme commercially viable for assessment organisations the cost per Level 1 candidate is likely to be in the region of **£450 - £500**.

11.2 Candidate Fees – Level 2

The estimated costs per candidate at Level 2 are likely to be the same as those for a Level 1 candidate, given that, if our recommendations are accepted, the diet of assessments is likely to be similar.

11.3 Candidate Fees – Level 3

The estimated costs of providing a Live Assessment Day at Level 3 are likely to be:

Venue Hire -	£700
Assessor's Fees -	£800 (based on £700 fee plus £100 travel expenses)
Expert's Fees -	£600 (based on £500 fee plus £100 travel expenses)
Administration -	£100

Therefore, the costs of running a Live Assessment Day at Level 3 would be in the region of £2,200. We anticipate that a maximum of 6 candidates could be processed in a day, which would equate to approximately £370 per candidate.

To that would have to be added the cost of assessing the portfolio and (providing detailed feedback where appropriate) which we put at £100.

The fee charged to candidates would also have to provide for developmental costs, administration costs and the costs involved in maintaining an external examiner system/assessment boards to review results. We estimate that the approximate cost per candidate in respect of these matters would be £150.

Therefore, the costs per candidate at Level 3, before application of any uplift to reflect profit margin, would be approximately £520. To make the scheme commercially viable for assessment organisations the cost per Level 3 candidate is likely to be in the region of **£575 - £625**.

11.4 Candidate Fees – Level 4

The estimated costs per candidate at Level 4 are likely to be the same as those for a Level 3 candidate, given that, if our recommendations are accepted, the diet of assessments is likely to be similar the diet of assessments is likely to be similar. However, Level 4 candidates will have to submit two pieces of judicial evaluation, which will need to be considered by the assessment organisation. We estimate that this is likely to add an additional cost of approximately £30 per candidate.

Therefore, the costs per candidate at Level 4, before application of any uplift to reflect profit margin, are likely to approximately £550. To make the scheme commercially viable for assessment organisations the cost per Level 4 candidate is likely to be in the region of **£600 - £650**.

11.5 Costs of any reaccreditation

If reaccreditation is part of the operational scheme it is anticipated that this would only be necessary after a minimum of 5 and a maximum of 10 years. **We would recommend any reaccreditation to require a reduced diet of assessment consisting probably of a cross examination exercise** which we estimate could be provided for £300 per candidate at Levels 1 and 2. Even if reaccreditation required, instead of cross examination, a portfolio, this could be properly marked and administered at Levels 1 and 2 for £160 per candidate For Levels 3 and 4 the enhanced nature of that cross examination assessment instrument would cost £450 per candidate. We estimate that a portfolio at Levels 3 and 4 could be assessed for £200 per candidate.

11.6 Costs of Regulation

We anticipate that, if responsibility for a QAA scheme passes to a regulatory/external body, then that body is likely to incur costs in respect of its governance of the scheme, including matters such as:

- receiving and dealing with applications;
- tracking evidence from applicants;
- tracking the status of applicants; and
- monitoring of standards of assessment organisations.

It is usual for some of the costs of regulation to be passed on to candidates; for example the SRA charges a fee per candidate in respect of its monitoring of the police station and duty solicitor accreditation schemes. The costs of regulation would, therefore, increase the overall cost to candidates of obtaining accreditation under a QAA scheme.

12 Summary and conclusions

The main aims of this project have been:

- to research, analyse and report on assessment options that can be used to effectively assess advocates against a defined competence framework;
- make recommendations to the QAA Project Team as to the most effective assessment route that best covers the 4 levels of advocacy to be tested in the pilot; and,
- to consider and make recommendations for any passporting or exemption from requirements for particular types of candidate.

The researchers did not define the levels or the competences (as set out in the competency framework). These were drawn up by work stream groups and signed off for testing by the Reference Group (a body consisting of practitioners, the judiciary and policy-makers from the representative bodies and regulators and the Legal Services Commission and Ministry of Justice).

In developing and testing assessment mechanisms, the Research Team has been assisted by a group of advisers (included a retired circuit judge, a Chief Crown Prosecutor, a clerk to the justices and a pool of current criminal practitioners from both main arms of the profession). Assessments were conducted by a team of assessors consisting of three staff from Cardiff Law School's Centre for Professional Legal Studies³¹ and seven experienced practising advocates (including two Queen's Counsel).

12.1 Numbers participating

The pilot aimed to conduct 250 assessments using a wide range of assessment mechanisms across the four levels. It included, for the first time in accreditation processes of this kind, judicially conducted evaluation of advocacy skills. The original scope of the pilot was extended beyond defence advocates to include up to 30 advocates from the CPS, taking our target for candidates to 280.

Candidates were volunteers and despite significant efforts by the research team and interested parties (including the judiciary) 110 candidates were assessed, of whom 22 were judicially evaluated and 98 produced other assessment data in time to be included within the evaluation.

This lower than hoped for level of participation has reduced our ability to fully test the assessment mechanisms, especially at the higher levels (Levels 3 and 4, where 14 and 7 candidates participated). Disappointingly, the number of judicial evaluations carried out has inhibited our ability to ascertain the value of that as a reliable assessment mechanism and to compare the other assessment

³¹ All former barristers or solicitors and experienced advocates and advocacy teachers, they included a former CPS grade 3 prosecutor and current recorder.

mechanisms against it to evaluate the extent to which assessor judgments on the quality of advocates were similar to those of judges.

Conversely, whilst the number of candidates participating in the pilot is low, we have been able to test the assessment mechanisms on a range of candidates from different occupational groups: 52 solicitors; 33 barristers in private practice; 7 CPS advocates and a small number (6) of FILEX and pupil barristers participated.

12.2 Choice of assessment instruments

Available assessment mechanisms were mapped against the competency framework and the definitions of the four levels to ensure that as full a range of competences as possible was covered at the requisite level. The assessment methods tested by the project were assessment of:

- the performance of an advocate in a real trial by the judge in the trial (**judicial evaluation**);
- a **portfolio** where a candidate reflects on their conduct of a real trial;
- an advocate's anonymised **written advocacy**;
- the performance of an advocate in three simulated hearings by an assessor (a legal **submission**, a **cross examination** and **evidence in chief** were conducted depending on the level at which the candidate was being assessed);
- the performance of an advocate in a simulated client **interview** (conference), to test competences directed specifically at client skills; and,
- a **multiple choice test (MCT)** of legal knowledge (covering substantive law, evidence, and procedure).

With the exception of judicial evaluation all assessment instruments were chosen to ensure that the results of any assessment could be verified from records of that assessment consistent with normal practice in any assessment regime and, in particular, to allow any future QAA scheme to be based on methods which would permit moderation, verification of, and appeals from, assessments.

Given the importance of the live skill of advocacy, and the considerable interest expressed from stakeholders in involving judges in the process, this expectation was relaxed for judicial evaluation.

For the simulations, candidates were provided with case documentation developed by the research team in consultation with the assessors. They were asked to prepare for the particular activities which they then performed in front of one or more assessors. Performances were recorded.

Multiple Choice Tests were conducted on a closed-book basis. To permit open-book Multiple Choice Tests would have changed and sometimes devalued the evidence gained from answers. Conversely the recognition that in reality an advocate might, without prejudice to the client, have been able to look up some answers resulted in our refining the data eventually derived from the Multiple

Choice Test to enable us to focus on the questions we would expect them to be able to answer 'on their feet' (the 'must-know' questions).

The portfolio enabled the candidates to produce a document which sets out certain specified aspects of cases they have dealt with, in order to satisfy stated criteria. The document enables an assessor to judge the extent to which the relevant competences are displayed and whether they are demonstrated at an appropriate level. The assessment criteria were formulated by the research team in consultation with the assessors. Portfolios were common to all Levels. In the written advocacy test candidates were asked to provide a suitably anonymised piece of written advocacy they had prepared for a real case. This provided an opportunity to assess the extent to which an advocate can research, construct and present a legal argument.

A larger range of assessment mechanisms was trialled at Levels 1 and 2 because of the well-founded expectation that the greatest number of volunteer candidates would come forward at these Levels, targeting resources at the areas where we would be most likely to get the most data. Furthermore, the assessment of the higher level skills means they may only be susceptible to assessment by a narrower range of mechanisms. The following table indicates the assessment mechanisms used at each level.

12.3 Assessment mechanisms by levels

Level	Judicial Evaluation	Portfolio	Written Advocacy	Cross examination	MCT	Interview	Submission	Examination in Chief
1	—	✓	✓	✓	✓	✓	✓	—
2	✓	✓	✓	✓	✓	—	—	✓
3	✓	✓	✓	✓	—	—	—	—
4	✓	✓	✓	—	—	—	—	—

Judicial Evaluation at Level 1 would have required the involvement of lay magistrates in the scheme and so was not attempted.

Whilst the above combinations of assessment mechanisms ensured that the vast majority of competences in the framework were covered, certain competences were not assessable through these mechanisms, in particular:

B3 Meets deadlines

- 1 Keeps the court informed of any timing problems/delays
- 2 Complies with judicially imposed timetables
- 3 Is punctual

C1.2 Keeps lay and professional client up-to-date; **and**,

C 2.3 Keeps lay and professional client up-to-date.

All of these competences deal with the handling of information about timings. It is our view that they cannot be assessed as part of an accreditation scheme without significant (and currently unfeasible) monitoring of practice.

We recommend that competences B3 (1, 2 and 3); C1.2 and C 2.3 be removed from the Competence Framework for the purposes of accreditation.

There are other competences within the framework which can only be subject to limited assessment by assessment mechanisms (such as C.2.1 Observes professional etiquette in relation to third parties), which essentially relies on the candidate self-certifying their competence in this regard.

We recommend that competence C 2.1 be removed from the Competence Framework for the purposes of accreditation.

12.4 Assignment of advocates to Levels

Designing a set of levels applicable to advocacy is a difficult task, but we have reservations about the current definitions which do not always clearly distinguish the types of case falling within each level. In particular, implementing these levels for the pilot revealed that candidates were often uncertain as to which level they should be assessed at. About half of candidates at Levels 1 and 2 were uncertain about their proper level based upon the cases they dealt with. These doubts were less marked at the higher levels, but were still present. Additionally the candidates were not always clear about the Level of the case they were using in their portfolio. For a scheme based on these Levels to be efficient when operational, it would be necessary for the advocate, their clerk (where relevant) and the paying body to be clear about the level of the case for which public funds were being sought. The fact that sexual offences are to be found at all three Crown Court Levels without clear guidance as to which offence falls into which level gives rise to particular uncertainty.

We recommend that the definition of levels is reconsidered prior to the implementation of any scheme and that particular consideration is given to defining sexual offences more specifically if the Levels as they currently stand are to work efficiently in an operational scheme.

The research team have devised an alternative way of identifying levels of case, set out at Appendix L, for consideration by relevant stakeholders.

12.5 Testing of simulated assessment mechanisms

Each assessment instrument at Levels 1 and 2 was pre-piloted on a sample of advocates not participating in the main pilot. During the pre-pilot, live assessments were video recorded and the resultant performances subjected to joint scrutiny by all assessors, who discussed the relative merits/demerits of what they saw and agreed the standard to be reached in each assessment. Through this process we were able to:

- ensure that the exercise allowed an assessor to make an evaluation based upon relevant competences;
- ensure that the exercise allowed an assessor to distinguish sufficiently between different levels of competence;
- check the completeness and accuracy of the guidance on points which should attract marks;
- ensure that the criteria and the layout of the mark grid allowed assessors accurately to record with sufficient particularity the aspects of a performance attracting marks, and those losing them;
- ensure that a basic level pass, and a fail at that level could be identified by reference to the use of the criteria and the attribution of marks;
- identify the characteristics of performances which should attract, or cause a reduction in, marks;
- ensure consistency of marking between assessors; and
- use the pre-pilot as an opportunity to train assessors in order to achieve greater consistency.

For Level 3 cross examination a slimmed down piloting of the process took place consistent with the involvement of two QC assessors and the limited number of Level 3 candidates. The assessors exchanged guidance on the assessments and discussed early performances whilst the Level 3 assessments were conducted. A further meeting took place with all three assessors once all Level 3 assessments had been completed, a selection of the recorded performances were viewed and the mark to be attributed was agreed upon to ensure consistency and an appropriate pass standard (of 60% on the agreed criteria). The assessors also formed the view that the assessment was capable of distinguishing between a Level 3 and a Level 4 performance (the mark which divided Level 4 advocates from those at Level 3 was 80%).

12.6 Judicial Evaluation

The judicial evaluation form was designed by one of the QC assessors and the principal investigator. A short set of explanatory notes was also developed and pre-piloted in Southwark Crown Court. It was not possible within project resources or time frames to conduct training of the judiciary in the use of the forms but

meetings were held at three courts³² to explain the pilot and our aim in enlisting judicial support. There was similarly no prospect of involving the judiciary in a process of moderation of assessments to ensure consistency. Consistency was to be assessed through statistical analysis of different judicial evaluations of the same candidates.

12.7 Results from the assessment mechanisms

It is important to recognise that those taking the pilot assessments voluntarily may not be typical of the professions as a whole. In particular, we do not know whether the levels of competence that the volunteer candidates demonstrated would be higher, lower or similar to those that would be expected under a mandatory accreditation scheme. Whilst one might expect that those participating as volunteers would be more confident than average of their competence, the fact that many candidates might use the exercise as a way of gearing up for the scheme and testing themselves at the level to which they aspired, might mean that their results would be poorer than would be expected should a scheme be rolled out beyond the pilot. As such, the levels of performance demonstrated by advocates under these assessment mechanisms should not be taken as being typical of standards of advocacy generally.

12.8 Level 1

Candidates generally passed Level 1 assessments. This may be consistent with the volunteer status of the candidates and the 'entry level' of accreditation status. Of 35 candidates who took Level 1 assessments:

- 4 (11%) failed the interview.
- 5 (14%) failed the submission.
- 2 (6%) failed the cross examination assessment. They both passed their interview. One had also failed on their submission.
- 12 out of 35 were found to be performing at Level 2 standards on the cross examination assessment.
- 3 (9%) who failed the interview reached Level 1 on cross examination and the other interview fail reached Level 2. 4 of the submission fails reached Level 1 (one did not).
- 8 (23%) Level 1 candidates failed the Multiple Choice Test.
- 2 out of 7 (29%) of those submitting portfolios at Level 1 failed. Both fails passed the interview assessment. 1 candidate passed their portfolio but failed the interview.

³² The research was expected to go forward in four court centres initially, but was extended beyond those four courts in an attempt to ensure a larger number of candidates volunteered.

- Only 4 candidates submitted written advocacy. There was one technical fail.

Analysis of the results comparing the different assessment mechanisms demonstrated a reasonably strong overlap between the interview and the submission assessments, suggesting that if other evidence points in the same direction it might justify relying on one or other of the assessments rather than both at Level 1.

12.9 Level 2

Pass rates for Level 2 candidates were noticeably lower. This may reflect the standard of the test (representing a sterner test of the fitness of candidates to work in the Crown Court) and the nature of volunteer candidates (who may have undergone the assessment to ascertain whether they might be sufficiently good to get a Level 2 accreditation in any future scheme).

- 20 out of 39 (51%) of those taking the cross examination assessment at Level 2 passed.
- 22 out of 39 (59%) passed examination in chief.
- 20 out of 39 (51%) passed the Multiple Choice Tests. Given concerns voiced by practitioners that Multiple Choice Tests tested information they could simply look up, we analysed separately the questions which one would expect them to know 'on their feet' (the 'must-know' questions).
- 21 out of 39 (54%) failed the must-know questions (got less than 4 out of 5 of these correct).
- 19 Level 2 candidates submitted portfolios and all passed.
- 20 submitted written advocacy assessments. There were two 2 fails and one technical fail.

Analysis of the results suggested there was a strong overlap between the assessments provided by cross examination and examination in chief. That is not to say that there is no variation between the two types of assessment, but if other evidence points in the same direction it might justify relying on one or other of the assessments. There is also some suggestion of a relationship between the written advocacy and examination in chief assessments, although this is based on a relationship between a very small number of fails.

12.10 Level 3

A limited number of candidates were assessed at Level 3 assessments:

- 10 out of 13 (77%) passed cross examination. 4 out of 13 were found to be performing at Level 4 standard by the assessors.
- 8 out of 10 (80%) candidates passed the portfolio.

- 7 out of 11 (64%) passed written advocacy.

Analysis of the marking mechanisms did not find any significant overlap between the mechanisms used.

12.11 Level 4

The number of candidates at Level 4 was low.

- 7 submitted portfolios (5 passed).
- 7 submitted written advocacy (4 passed).

One candidate failed both assessments.

12.12 Judicial Evaluation

Judicial Evaluation of advocates has the significant advantage of being based on assessment of actual live performance by highly experienced and respected members of the criminal justice system. It is also supported by key stakeholder groups in the criminal justice system. Such support has been expressed subject to reservations, in particular:

- The potential for a judicial role as an assessor to compromise their independence;
- The potential for assessments to be subject to disclosure as part of an appeal by a criminal defendant against conviction and by a candidate seeking to appeal a failed application for accreditation;
- The potential for disclosure to militate against candour, particularly when faced with poor performance;
- The additional workload Judicial Evaluation would place on the judiciary (and any body administering a scheme of Judicial Evaluation);
- The need to ensure candidates are assessed in respect of multiple cases (to ensure they are assessed on a full range of competences and are not assessed on the basis of idiosyncratic cases); and
- The need to monitor Judicial Evaluations for consistency and fairness between candidates (to counter any perceptions of bias).

It is not the function of this research to resolve the debates around independence and the confidentiality of the process. This research sought to begin the process of testing the practicality of Judicial Evaluation and compare the results from such evaluations with other assessment mechanisms (thereby acting as a test of the overall robustness of different assessment mechanisms, including Judicial Evaluation itself).

As noted above, Judicial Evaluation was sought at Levels 2, 3 and 4. Advocates indicated courts in which they regularly appeared, or expected to appear over the

pilot period. Those courts were contacted with a list of the names of candidates who might appear in their court. The LSC QAA team followed up on this contact to encourage evaluations to be forthcoming. We hoped for one evaluation of each candidate at Levels 2 and 3, with 2 for each at Level 4.

The questionnaire was to be used sufficiently contemporaneously to provide an accurate reflection of performance. It was also developed to be used in the absence of detailed training or briefing, being sufficiently concise to encourage maximum levels of participation from the judiciary.

The number of candidates who could have been judicially evaluated under the pilot was 148. Of these we received Judicial Evaluation for 22. We received more than one evaluation for only three candidates. It was possible that some of the remaining 126 candidates had not conducted trials during the relevant period, we sought data on the trials that candidates had conducted during the relevant period. 44 responded to our requests. Between them they had conducted 137 trials but we received only 9 Judicial Evaluations from those trials. Whilst not all of these trials would have been assessable (some may not have been heard by full time judges), the low level of response demonstrates the difficulties which may attend asking judges to provide evaluations as part of an operational scheme. Some of the failure to respond may be due to administrative problems or workloads for the judges. Some judges were opposed to the involvement of the judiciary in the pilot, or the broader process of QAA, and we understand did not participate on that basis. Similarly, we are aware of judges who did not assess candidates whom they regarded as performing in the trial in question at below the normal standards of that advocate.

Were the low levels of response we experienced to be overcome, it would require significant investment in an administrative mechanism, and support for judges designed to assuage any concerns they have. Candidates would be likely to have to wait some time before evidence for assessment was available and, the candidate might have to be assessed on types of hearing other than trials, which would not enable assessment in all the competences.

12.13 Results for judicial evaluation of all candidates

For the 22 candidates who were judicially evaluated, only 10 had completed the other assessment mechanisms, hampering our ability to compare the results of judicial evaluation with the other assessment mechanisms.

As we have only three candidates for whom we can compare the assessments of different judges regarding the same candidate, we do not have sufficient data to enable testing of the reliability of judicial evaluation as part of an accreditation scheme. We therefore include the comparisons between the three candidates for interest only. The judges were in agreement about one candidate. For the second candidate, one judge rated them one point higher on the scales than the other (the difference between a 4 and a 5). For the third the variation was between one and

two points on the scales (i.e. assessment on the various criteria range between a 3 and a 5).

12.14 Judicial Evaluation in an operational scheme

Our data gives significant cause for concern about the practicalities of instituting a scheme of Judicial Evaluation that contributes meaningfully to QAA.

We cannot therefore recommend Judicial Evaluation as an assessment mechanism for the majority of advocates.

Judicial Evaluation would require the need for training in the grading of aspects of advocacy performance; the independent recording of the performance [transcripts]; the provision for appeals consistent with other mechanisms; and the need to ensure that Judicial Evaluation is based on data from a number of the advocate's cases to ensure a full range of competences was tested, that assessments were not prejudiced by idiosyncratic cases/clients and to enable some monitoring of consistency of approach across the judiciary.

If Judicial Evaluation is to be developed as part of a QAA process **we believe the practical difficulties should be addressed on a cohort of candidates where the numbers are not overwhelming in practical terms and where judicial commitment to the exercise is likely to be at its highest. This suggests concentrating on Level 4 assessments, where there is the strongest prospect, initially, of developing judicial evaluation into a valuable tool.** Given that judicial responses may be more forthcoming for candidates with whom they have existing experience, and the potential reluctance of judges to grade performances which are poor, any such scheme is more likely to resemble a system of judicial references. The utility and practicality of reference-based approaches gives rise to its own concerns.

12.15 Which assessments for which level? Recommendations for the implemented scheme

The research team, assessors and our own feedback from candidates supported the view that the cross examination assessment was the most important competence assessment. It has a number of advantages: it assesses a wide range of competence; in a context as similar to genuine advocacy as possible and on a basis which is fair and replicable (candidates sit the same standard of test); and it is not prone to problems that portfolios and other assessment tools are prone to (that they test a candidate's presentational skills as much as they test their actual skills). A further advantage of a simulated cross examination is that it can be used if required as a gatekeeper to assisting the progress of candidates up the levels.

Whilst portfolios can be managed by candidates in order to present themselves in their best light (problems more particularly rehearsed at section 2.3.2 in the report), the team of assessors felt that they provided good insight into the level of experience a practitioner has achieved and their ability to reflect on that experience. Interestingly, and to the surprise of assessors, candidates who

undertook these were appreciative of the opportunity to reflect upon and relate their own practice.

One aim of the research and evaluation conducted was to reduce the number of assessments to the minimum which could properly assess the competences. Interviews (or conferences) and submissions were shown by the data to give results which overlapped significantly, suggesting that only one of these assessments is needed in an operational scheme. It should be noted however, that there are aspects of the competences – concerned with relating to the client – which can only be assessed live through interview, but it would be possible to judge these in part from a portfolio.

There is one competence, **A.1.3** (*Making only relevant submissions*), which we consider capable of assessment in the recommended regime of assessment for Levels 2, 3 and 4 only through Written Advocacy. Written Advocacy does not appear as an assessment instrument in the recommended diet at Level 1, where the research and feedback indicates that it is not appropriate. Thus if this competence is to be assessed at Level 1, **we recommend that the Submission advocacy assessment be retained at Level 1**. The alternative would be to dispense with that competence at Level 1 and thereby the need for the Submission advocacy assessment. **The interview can, we suggest, reasonably be dispensed with** because of the correlation in the data between it and the submission and because other relevant areas of competence can be addressed in the portfolio.

Assessment Recommendations

As a result of this pilot, we believe that an assessment landscape which balances the need to assess a range of competencies with the proportionate testing of those competences can lead to a simpler assessment regime than that trialled in the pilot:

Level 1

We recommend:

- A portfolio to include 3 trials at Level 1; at least 2 to be of either way offences;
- A submission based piece of advocacy;
- A simulated cross examination assessment, and
- Specifically if **D 3.3** (*Assists the court with the proper administration of justice*) above remains a competence, then **either** a Multiple Choice Test with a relevant question or a requirement in a portfolio case to show that competence would be needed. However, since **B.2.1** (*Complies with appropriate Procedural Rules and judicial directions*) remains at this Level an untested competence, a preference for the Multiple Choice Test option would enable the assessment of both competences.

No written advocacy is recommended at this level due to the fact that our cohort was rarely called upon to do this. They could not provide such evidence in any

realistic way. The need for this could be kept under review should practice change in magistrates' courts.

Level 2

We recommend:

- A portfolio to include 2 trials at Level 2 and a piece of Written Advocacy;
- A simulated cross examination assessment; and,
- depending on the view taken about capture of competences D 3.3 and B 2.1 in relation to Level 1, a Multiple Choice Test.

Level 3

We recommend:

- A portfolio to include a single Level 3 trial and a piece of Written Advocacy;
- A simulated cross examination assessment.

Level 4

We recommend:

- A portfolio to include a single Level 4 trial and a piece of Written Advocacy;
- A simulated cross examination assessment;
- At least 3 pieces of Judicial Evaluation. Two of these on Judicial Evaluation forms, one covering the same case as in the portfolio (the advocate would have to alert the assessment organisation so that the court could request that the judge do the evaluation); a second from another case, and a third a narrative document detailing the advocate's strengths and weaknesses and provided at the request of the advocate by a judge (perhaps of High Court level) who has been the judge in a trial lasting at least 3 weeks in which that advocate appeared.

If competence **D.1.3** (*Maintains pace throughout the course of the trial*) were modified by the deletion of the words in italics, it could be assessed in the above regime. If not it would remain unassessed.

If such a scheme were to become operational it would leave only 3 of the competences not already recommended for deletion with any uncertainty over their being assessed. These are:

- **C.3.1** Observes restrictions and judicial rulings on questioning
- **D.3.1** Observes duty to the court and duty to act with independence
- **D.3.2** Advises the court of adverse authorities and, where they arise, procedural irregularities

Each could be satisfied by a requirement (in the guidance to candidates on portfolios) for them to choose cases which showed particular challenges on these points.

Ensuring flexibility and facilitating transition between the levels

A danger of any level-based system of accreditation is that it is operated inflexibly and prevents capable advocates developing the skills and caseloads necessary to practise competently at the higher levels.

Our recommendation is that any scheme should have ways of testing some competences that assist this transition to a higher level. A candidate could then pass this transitional or gatekeeper assessment and be permitted to take cases at the higher level for a probationary period (of say 12 months) whilst they completed the full diet of assessments for that higher level.

The assessors testing our simulated cross examinations have indicated their confidence that the cross examinations set for Level 1 and Level 3 candidates can be used to test whether candidates at this level are also performing at higher levels of competence (Level 2 and Level 4).

We recommend that the simulated cross examination assessments be used as this gatekeeper assessment for transition to the higher level.

If the recommendation is accepted, **there would have to be Registers of Probationers at each level to ensure proper monitoring of the QAA scheme and for the proper attribution of cost to cases.** A body would need to take responsibility for managing such a register. The LSC currently has responsibility for similar registers for police station and court duty solicitors through First Assist.

Reaccreditation

Any QAA scheme should cover the timing and process for reaccreditation. There is a need to ensure that reaccreditation processes pick up changes in a candidate's competence over time, balanced with the need to ensure candidates are not affected by a disproportionate assessment regime.

We recommend that a new QAA scheme should not require any reaccreditation within a period of 5 years from the time of an advocate's first accreditation, though there may be some support for having a lapse of 7 years. In the shorter time frame many advocates in the earlier part of their career will in any event have sought accreditation at the next level. If they have not it would appear prudent to reassess. **Reaccreditation might be by way of a reduced diet of assessments.**

Whilst a portfolio is the cheaper mechanism to assess, it does not have the rigour of a live assessment. An adequately prepared portfolio is also likely to be more costly in terms of a candidate's preparation time. **We recommend that reaccreditation is by way of a simulated cross examination.**

12.16 Passporting and associated issues

The team was asked to consider whether and where it would be possible to make recommendations for a reduction in the number or need for assessment of any

group of candidates based upon their previously acquired status or qualifications. We make specific recommendations regarding these matters below. In addition to those recommendations,

we recommend that appropriate exemptions (which could include exemption from all relevant assessments) be given to participants based upon successful performance in the pilot. This is in addition to any recommendations which we make in respect of the particular professional groupings discussed below.

Analysis of variations in assessment scores related to occupational grouping is relevant to any consideration of passporting. Differences in quality might be expected to occur given the different professional backgrounds of the candidates in terms of the extent to which they were trained in, and have subsequently specialised in, advocacy at the level being assessed.

The following tables summarise the results under each assessment by professional grouping. We have done this for levels 1 and 2 where there are a reasonable number of assessments. Whilst the tables may be read as suggesting some differences, none of these differences appear to reach or near statistical levels of significance.³³

Table 19: Percentage Failure Rate of Compulsory Level 1 Assessments by Professional Grouping (Level 1)

	Professional Grouping					
	Barrister	CPS	FILEX	Pupil	Solicitor	Total
Interview	14.3%	0.0%	33.3%	0.0%	10.0%	11.4%
Submission	14.3%	0.0%	33.3%	33.3%	10.0%	14.3%
Cross examination	0.0%	0.0%	0.0%	33.3%	5.0%	5.7%
MCT	42.9%	0.0%	33.3%	66.7%	10.0%	22.9%
Must-know MCTs	28.6%	50.0%	100.0%	66.7%	35.0%	42.9%
N	7	2	3	3	20	35

³³ Chi-square tests p for all comparisons all > .1

Table 20: Percentage Failure Rate of Compulsory Level 2 Assessments by Professional Grouping (Level 2)

	Barrister	CPS	Solicitor	Total
MCT	25.0%	40.0%	57.7%	48.7%
Examination in Chief	37.5%	80.0%	38.5%	43.6%
Cross examination	37.5%	80.0%	46.2%	48.7%
Must-know MCTs	25.0%	60.0%	50.0%	46.2%
N	8	5	26	39

The relatively low number of candidates assessed makes it hard to recommend any passporting in respect of already held qualifications. The differences between the different occupational groupings were not statistically significant and there does not thus appear to be a justification for permitting passporting on the basis of any particular occupational grouping, particularly given the level of failures across each of the groups (which may relate to their volunteer status). We have also analysed the results of candidates with external CPS gradings and the results do not suggest a basis for using those as part of a system of passporting or exemption.

However **we do make recommendations where there is other information which we believe leads to a defensible approach to exemption, with the proviso that the scheme be subject to review.**

There is no evidence to show that Duty Solicitors performed any better at Level 1 than solicitors who were not Duty Solicitors. However the Level 1 cohort of 20 solicitors consists of only 4 who are not current Duty Solicitors – and one of them was previously a Clerk to the Magistrates and Duty Solicitor. The vast majority of that cohort passed the assessments using instruments which we recommend as part of an operational scheme. One failed the cross examination and 2 the Multiple Choice Test.

It could be argued that this cohort is likely, given that the majority are Duty Solicitors, to include those with significant experience as advocates in magistrates' courts. They have already been scrutinised in order to achieve that Duty Solicitor status. The assessment instruments used for the purposes of Duty Solicitor accreditation lack any assessment of witness handling or of knowledge (as in the Multiple Choice Test). Nonetheless this cohort has shown itself largely competent on those tests. This may be because of the exposure they have gained. It should also be noted that Level 1 candidates were volunteers and there are plausible reasons for speculating that they might appear to be of a higher quality than a random sample of candidates.

A suitable balance should be struck between recognising their prior accreditation and assessing their quality. We therefore recommend that **Duty Solicitors should be able to acquire Level 1 status by passing only the cross examination assessment.** Other solicitors would need to take the full diet of assessments recommended below.

Another matter on which we would make a recommendation is in respect of all solicitors for whom criminal practice forms the main part of their work, and has done for a considerable period. **It may be that, in recognition of this, any future scheme could trial giving exemptions on all but the cross examination exercise to any solicitor with a criminal practice with a minimum number of years' (say 10) experience in this field.**

Barristers

Eight barristers who were not pupils were assessed at Level 1 (2 of them work for the CPS). Three barristers failed the Multiple Choice Test, but none failed the cross examination. Both CPS candidates passed all assessments. Only 1 barrister failed more than one assessment. The numbers are insufficient to make wholesale recommendations based solely on this data as to passporting or exemptions. It should however be remembered that this cohort had on average only 1.9 years' criminal practice. **We recommend consideration be given to allowing a barrister exemption from any cross examination and submission test at Level 1..** This would be an appropriate recognition of the training and assessment they have undergone on the BVC. To do this would also sit most consistently with the findings about barristers at Level 2.

If a Multiple Choice Test were part of the diet of assessments there is no basis upon which to suggest any exemption be offered in respect of this assessment.

We therefore recommend that the Multiple Choice Test be used as the gatekeeper assessment for barristers at this Level. Passing the Multiple Choice Test would trigger a 12 month probationary period within which a portfolio must be passed.

Level 2

What the evidence from the Level 2 assessments suggests with some force is that there is no evidence to show that merely having the right to appear and conduct trials in the Crown Court means that an advocate can exercise their skills at the requisite Level for that court.

We recommend trialling a full diet of assessments for entry to this Level coupled with a reduced diet for those already working at this Level, with a minimum requirement of a cross examination assessment for all to be passed within a year of the inception of any operational scheme.

Level 3

Numbers assessed at this Level were too small to be able to make recommendations for passporting. **We suggest that any future scheme is based on the requirement that a new entrant at this Level take a full diet of assessments, and for those whose practice is already at this Level the requirement to have passed that diet of assessments within 2 years of the inception of an operational scheme.** The greater length of time takes account of the fact that those practising at this level will have fewer and longer trials.

Level 4

Numbers assessed at this level were too small for us to make specific recommendations regarding passporting.

Our development and testing of the Level 3 cross examination and feedback from Level 4 candidates who would have welcomed the opportunity to conduct a live assessment, leads us to believe that such assessments can be designed and implemented at proportionate cost.

We therefore recommend the addition to the diet of assessment piloted at Level 4 a cross examination exercise of the same degree of complexity as that taken by Level 3 candidates, but assessed to the more exacting standards expected of a Level 4 candidate. This exercise affords the possibility of achieving a pass at Level 3 or 4 for Level 3 candidates, providing an indication of their readiness to proceed to other Level 4 assessments and facilitating progression through the scheme.

We suggest that any future scheme is based on the requirement that a new entrant at this Level take a full diet of assessments, and for those whose practice is already at this level, the requirement to have passed that diet of assessments within 2 years of the inception of an operational scheme. The greater length of time takes account of the fact that those practising at this level will have fewer and longer trials.

12.17 Equalities and diversity

There was inadequate data to compare the performance of candidates to test for diversity concerns. **This is a matter which we recommend is monitored as part of the process of implementation.**

Assessors were trained to recognise the need to consider potential equality and diversity issues if they arose in portfolios. Any problems of this sort would be marked under the criterion relating to ethics.

12.18 Setting up an operational scheme

The collaboration of funders, the professions and the judiciary is an important element in the establishment of a QAA scheme which carries sufficient support amongst those participating in the scheme (as candidates or assessors). **Our recommendation is that this work is overseen by a joint body of funders, the judiciary and the professions (an Advocacy Board) which would have responsibility for the establishment of a scheme, regulations, validation and periodic monitoring of assessment organisations who would run assessment diets for which they were authorised.**

The experience of the pilot is that successful and consistent assessment relies on a variety of skills which are not to be found solely in either those working in education or current practitioners. Considerable educational experience is required for the detailed work of preparation of assessments, and guidance in using

assessment criteria, as well as for the operation of fair assessment procedures which needs to be matched by a constant dialogue with current practitioners about standards, practices and practicalities of the skills employed in assessment exercises.

It is also important, in our view, that assessment organisations have adequate administrative support from an institution experienced in assessment.

12.19 Feedback to candidates, appeals, monitoring and ensuring consistency across assessment organisations

We recommend, consistent with current practice, that candidates are provided with feedback in their performance in appropriate terms under each of the assessment mechanisms used. Special consideration may need to be given to whether such requirements can be met in relation to judicial evaluation, if that is proceeded with. Any scheme will also need to make provision for appeals against assessments consistent with current approaches operating in other assessment schemes.

The greater the number of assessment organisations validated, the greater the need to monitor comparability of results across assessment organisations. If, as is to be expected, there is a commercial market for assessments, the risk of forum shopping by candidates to purchase the easiest assessment needs to be avoided.

We recommend adequate reporting of assessments (e.g. pass rates), with differences in pass rates capable of further investigation. The advocacy board should also consider what further monitoring of assessment organisations is required through scrutiny of assessment instruments, requirements for double-blind marking and records of assessment decisions, for example. A second way to ensure continued fairness would consist of annual joint meetings of representatives of the markers at each level from all authorised assessment organisations, where sample assessments could be compared and guidance given as to where the appropriate standard should lie.

12.20 Costs of an Operational Scheme

The costs of obtaining accreditation will be a significant concern for practitioners and this is reflected in the Commission's summary of candidate feedback (para. 2.9.1). Therefore, we have endeavoured to estimate the likely cost to candidates of completing accreditation at each of the Levels. The costings draw on both our experience of running the pilot assessments, and our experience of running schemes which include a similar diet of assessments, such as CLAS (Criminal Litigation Accreditation Scheme).

Any estimate of costs has to make provision for:

- developmental costs (particularly preparation of materials);
- administration costs;
- overheads (such as venue charges);

- assessors' fees (including live assessment days, marking of portfolios/written advocacy and attendance at Test Boards);
- actors' fees and fees for expert witness for cross examination assessments at higher Levels;
- fees for external examiners.

Our estimates focus on the cost of the assessments themselves and do not include costings for the setting up and administration of the scheme as a whole.

If the QAA scheme is to be offered via assessment organisations, then those organisations will wish to secure a financial return, so an uplift to reflect profit margin would also need to be applied. The amount of that uplift will vary depending on the nature of the organisation, the extent to which it is able to make any economies of scale, the administration system already in place and the extent to which it can tie in the development of this scheme with any others it has run.

The amount to be charged will also vary depending on whether or not the organisation has to charge VAT. All costs and likely fees are quoted exclusive of VAT and are conservatively drawn – taking no account of the factors above which might operate to reduce either the suggested costings or the impact of the uplift applied in respect of profit.

It should be remembered that some of the advocates already pay significant amounts for other forms of professional accreditation. If a new scheme replaces, wholly or partially, any such accreditation there would be either total or partial saving of those fees.

Additionally it would be possible for the professional regulatory bodies to recognise any accreditation - or reaccreditation - under QAA for the purposes of the requirements of compulsory professional development. This would both ensure that CPD undertaken was entirely relevant to the lawyer's area of practice and would also have the effect of reducing the real cost to an individual advocate of QAA assessment.

Candidate Fees – Level 1

We estimate the average costs of providing a Live Assessment Day at Level 1 as:

Venue Hire -	£700
Assessor's Fees -	£700 (based on £600 fee plus £100 travel expenses)
Actor's Fees -	£250 (based on £150 fee plus £100 travel expenses)
Administration -	£100

Therefore, the costs of running a Live Assessment Day at Level 1 would be in the region of £1,750. We anticipate that a maximum of 8 candidates could be processed in a day by such a team. With such numbers the figure per candidate would therefore be £220.

To that would have to be added the costs of assessing a portfolio (providing detailed feedback where appropriate) and marking the multiple choice test which for this purpose we set at £60 and £10 respectively.

Additionally, the fee charged to candidates would also have to provide for developmental costs, administration costs and the costs involved in maintaining an external examiner system/assessment boards to review results. By way of guidance we would expect each single set of case study papers as well as an MCT to cost (including developmental meetings) in the region of £2,000. Case studies and MCTs have to be used cyclically and revised. They can therefore only be used a limited number of times. We estimate that the approximate cost per candidate in respect of all the above would be £110.

Therefore, the costs per candidate at Level 1, before application of any uplift to reflect profit margin, would be approximately £400. To make the scheme commercially viable for assessment organisations the cost per Level 1 candidate is likely to be in the region of **£450 - £500**.

Candidate Fees – Level 2

The estimated costs per candidate at Level 2 are likely to be the same as those for a Level 1 candidate, given that, if our recommendations are accepted, the diet of assessments is likely to be similar.

Candidate Fees – Level 3

The estimated costs of providing a Live Assessment Day at Level 3 are likely to be:

Venue Hire -	£700
Assessor's Fees -	£800 (based on £700 fee plus £100 travel expenses)
Expert's Fees -	£600 (based on £500 fee plus £100 travel expenses)
Administration -	£100

Therefore, the costs of running a Live Assessment Day at Level 3 would be in the region of £2,200. We anticipate that a maximum of 6 candidates could be processed in a day, which would equate to approximately £370 per candidate.

To that would have to be added the cost of assessing the portfolio and (providing detailed feedback where appropriate) which we put at £100.

The fee charged to candidates would also have to provide for developmental costs, administration costs and the costs involved in maintaining an external examiner system/assessment boards to review results. We estimate that the approximate cost per candidate in respect of these matters would be £150.

Therefore, the costs per candidate at Level 3, before application of any uplift to reflect profit margin, would be approximately £520. To make the scheme commercially viable for assessment organisations the cost per Level 3 candidate is likely to be in the region of **£575 - £625**.

Candidate Fees – Level 4

The estimated costs per candidate at Level 4 are likely to be the same as those for a Level 3 candidate, given that, if our recommendations are accepted, the diet of assessments is likely to be similar. However, Level 4 candidates will have to submit three pieces of judicial evaluation, which will need to be considered by the assessment organisation. We estimate that this is likely to add an additional cost of approximately £30 per candidate.

Therefore, the costs per candidate at Level 4, before application of any uplift to reflect profit margin, are likely to be approximately £550. To make the scheme commercially viable for assessment organisations the cost per Level 4 candidate is likely to be in the region of **£600 - £650**.

Costs of any reaccreditation

If reaccreditation is part of the operational scheme it is anticipated that this would only be necessary after a minimum of 5 and a maximum of 10 years. **We recommend any reaccreditation to require a reduced diet of assessment consisting of a cross examination exercise** which we estimate could be provided for £300 per candidate at Levels 1 and 2. If reaccreditation required a portfolio rather than a cross examination assessment, this could be properly marked and administered at Levels 1 and 2 for £160 per candidate. For Levels 3 and 4 the enhanced nature of that cross examination assessment instrument would increase the cost to £450 per candidate. We estimate that a portfolio at Levels 3 and 4 could be assessed for £200 per candidate.

Costs of Regulation

If responsibility for a QAA scheme passes to a regulatory/external body, then that body is likely to incur costs in respect of its governance of the scheme, which will need in whole or in part to be passed on to the candidate.