UNIVERSITY of CAMBRIDGE
International Examinations

External Evaluation of the European Baccalaureate

Contract Notice No. 2007/S239-290368

January 2009

FINAL REPORT

Table of Contents

Executive Summary

Introduction

In 2007 the Board of Governors of the European Schools agreed to carry out an external evaluation of the European Baccalaureate, the objectives of which were:

- To determine to what extent the European Baccalaureate is fit for purpose
- To obtain recommendations regarding the measures to be taken to guarantee the quality of the European Baccalaureate and its recognition by the member States
- To identify the questions to be taken into consideration, and the actions to be undertaken, in order to be able to offer the European Baccalaureate to more students outside the European Schools.

The External Evaluation was conducted between May – December 2008.

Methodology

In answering questions relating to the continuing 'fitness for purpose' of the qualification, the key questions of the Evaluation related to an assessment model which examines four inter-related dimensions: validity, reliability, impact and practicality.

The Evaluation reviewed a range of primary and secondary sources, complemented by judgemental analysis, surveys, interviews and observations of process.

A particular focus was given to a comparability study looking at five subjects: English as Language 1, French as Language 2, Mathematics, Geography, and Biology.

The review of standards between different Examination Boards involved a project methodology which compared written papers, mark schemes and scripts in the designated subjects with equivalent qualifications in Germany, Ireland, France, England, Finland and Sweden.

For each element of a qualification being compared (whether within the European Baccalaureate or from the European Baccalaureate across to other qualifications) the comparison where possible was based on one expert judge analysing 5 scripts in each subject in the top 20%; 5 in the modal point of the mark distribution, 5 above pass and 5 below pass in each qualification.

The analysis used in the comparability work was one of cross-moderation. This approach was recommended as the most effective for the Evaluation given the limitations of time, the over-bearing scale of trans-national paired comparison methods and the lack of availability of sufficient numbers of judges, familiar with the European Baccalaureate and the National comparator.

The Evaluation also referred to a number of secondary sources of standards data available in equivalences linked to the currency and recognition of the European

Baccalaureate. We complemented this with a secondary analysis of further studies and material such as PISA analyses, and the Analysis of the Academic and Professional Careers of Graduates of European Schools, published in October 2008.

Comparability outcomes

In terms of the findings of the comparability study, outcomes reveal a very mixed picture across the different European Baccalaureate subjects – there is no common relationship, such as all subjects being broadly aligned in treatment with other countries, or a consistent relation emerging in terms of demand and standards. None of the analysts identified incoherence or grossly inappropriate content, approaches, or demand in the European Baccalaureate; what emerged were differences linked principally to variations in purpose and commitments. There is one subject, Geography, which appears to require urgent review.

The curriculum

While the formal taught curriculum in years 4-7 is a sound preparation for further academic study, its diversity of language choices and strong Science provision is offset by a relatively restricted range of subjects. Business-related and applied subjects, non-European languages, drama and media studies are all subjects of increasing popularity, both in schools and universities. The establishment of a more formal relationship with the Common European Framework of Reference for Languages offers both opportunities for harmonisation across languages and increases the transparency of students' attainment and provides a better preparation for the world of work. Science syllabuses themselves should be updated and a stronger and more coherent approach to the development of enquiry-based and investigative skills established. Cross-curriculum options are increasingly made available as a platform for the development of personal research and presentation skills. The *Travaux Personnels Encadrés* in France provide a model of this. That European Schools have developed great expertise in the education of students for whom the language of instruction is not the mother tongue is apparent. On a small sample, provided by the Report of the 2008 European Baccalaureate, better than average performance in a number of key subjects is shown and overall, the educational context of the European Schools has potential to demonstrate the beneficial impact of Content and Language Integrated Learning (CLIL) and to make a contribution to its development.

Assessment review of the European Baccalaureate

The European Baccalaureate involves a high volume of internal assessment by teachers. This is a potential strength, establishes an integrated learning and assessment model and makes a positive contribution to its validity. However, while European Baccalaureate teachers are very experienced, opportunities for ensuring that all teachers have access to early induction and standards training are vital. A survey of teachers, conducted as part of the Evaluation, indicated a low-frequency for such training which creates a potential threat to the reliability of the qualification.

The extent to which common standards can be shown to apply across all subjects is also an issue in the marking of final examinations where systems of marking review across subjects, between examiners and across years are not well-defined. To make this point is not to claim that such marking is not reliable, but rather to say that there is no mechanism by which such reliability can be demonstrated.

Examination entries, for the most part, are small in volume and in consequence do not lend themselves to strong statistical analysis. The establishment of in-year and year-on-year cross-moderation judgemental analysis would provide the Board of Governors with confidence that a profile of different results in different subjects or in different schools is attributable to a genuine difference in performance rather than an absence of marker reliability.

Qualification outcomes must be clearly read by those who use them. Existing definitions within the European Baccalaureate express subject 'size' in terms of the number of weekly periods which they entail. Greater clarity would be achieved by statements of actual time involved by the programme. In terms of weightings between different parts of the programme, the Evaluation was of the view that the value contributed by the internal assessment of Preliminary Marks should be retained. Proposals for a revised weighting of written examinations relative to L1 and L2 oral examinations would seem to overstate the contribution which a student's oral performance in languages makes to their overall European Baccalaureate score, particularly for those students who are preparing for science, medicine and engineering courses at university.

In terms of the setting of European Baccalaureate examinations, the practice of asking teachers to provide draft papers for examinations presents a number of risks to the system: incomplete sampling of the syllabus, predictability and a resistance to curriculum revision. It is possible to provide tools and training to avoid such risks.

The practice of double marking should be reviewed. Assessment research tends to demonstrate that teachers are well placed to produce a rank order of their own candidates even when the marks they award may not be in line with an examination standard. A revision of the practice of averaging the teacher's and the examiner's mark might be replaced by one in which the teacher provides a rank order for comparison against which that produced by a standardised examiner.

Currently the marking of final examinations is carried out on a residential basis. There is no particular advantage in the co-location of all examiners. Opportunities for the discussion of standards can be delivered in alternative ways – and many of the examination boards in England are moving towards 'virtual' standardisation approaches, particularly using digitised scripts and on-line marking.

Oral assessments play an important role in the Baccalaureate. The Evaluation proposes a number of opportunities for improving quality. The development of assessment criteria appears well in hand but improvements may be made in timetabling, the central construction of assessment tasks should be centralised and duration of some extended.

In considering the roles and responsibilities of key stakeholders involved in the European Baccalaureate, the involvement of the Board of Inspectors in curriculum

development and its implementation plays a vital role. In terms of the quality assurance of the European Baccalaureate examinations, however, their role might be more effectively defined away from direct scrutiny of the system, towards a more co-ordinated curriculum development and implementation.

Levels of university recognition

In reviewing the levels of recognition that Higher Education accords to the European Baccalaureate, a particular focus was given to universities in the UK and in France. The outcomes of the *Analysis of the Academic and Professional Careers of Graduates of the European Schools* provided a detailed source of information about the strength of the European Baccalaureate's predictive validity which is greater than that achieved by national qualifications such as A Level in England. The data of this report should form the basis of awareness raising websites and guidance linked to university admission. Ten universities in the UK who are all members of the prestigious Russell Group were surveyed for admissions policy. Information of use to applicants and to admissions tutors was very variable but there were positive instances of recognition and enthusiasm, particularly in the case of Oxford University.

In respect of admission policy in France, it is clear that while Article 5 (2) of the Statute of the European Schools provides admission to university, awareness levels were very low in those departments such as medicine and law likely to have competitive entry.

Expansion models

The expansion of the European Baccalaureate formed one of the drivers behind the commissioning of this Evaluation and we have looked at a number of models on which expansion might be based. The first, Model A, would involve the adoption of the European Schools' curriculum in its entirety. The proposed re-establishment of Culham as a European Academy provides an example of how this might work. Absorption into the state sector provides a funding stream and enables admission to the school to remain free for students. However, inclusion within the state sector also requires compliance with national educational policy that might in turn impact strongly on the mission of the school. Model B entails adoption of the European Baccalaureate for students in years 6 and 7 alone in a way similar to the IB Diploma. The number of language sections, in line with the recommendations of the Troika Working Group[1], would as a minimum involve a single language section and a restricted (vehicular languages only) range of L2, L3 and L4 choice. The recognition of multilingualism and the European Schools' acknowledged expertise in this area has the potential to offer significant opportunities for adoption of this model but a third approach, Model C, designed around the 'core' of the European Baccalaureate might also be considered. This is likely to be attractive in many countries, and to many students, as a way of demonstrating high levels of language performance and links to the CLIL agenda. In the context of all of the expansion models we consider that more

[1] Report of the Troika Working Group II – European Baccalaureate and Cooperation with other schools, Ref.: 2005-D-342-en-4,

explicit links with the Common European Framework for Languages is likely to prove both necessary and beneficial.

New technology

A further term of reference of the Evaluation was to consider the extent to which new technology would offer solutions to the issues experienced by the European Schools in the logistical and operational processes involved in paper-based examinations.

In respect of on-line test taking, the Evaluation suggests that this is piloted in respect of students' Preliminary marks. We consider it unwise to introduce on-line assessment into the final examinations without first having piloted its effectiveness. Not all subjects are suitable for on-line assessment; a similar level of facilities is not available to all students and security requirements tend to be more exacting even than with paper-based examinations.

In respect of marking on-line, however, the technology is increasingly becoming established to enable this. We suggest that pilots are carried out to identify feasibility but suspect that proposals for schools to scan their own students' scripts will not prove entirely effective.

Management of costs

In looking at the management of costs in the European Baccalaureate, the period between 2006 and 2008 saw a 3% increase in candidate numbers. Over the same period, operating costs have risen by 17.6%. In the broadest of terms, the costs of the European Baccalaureate run at a level three times greater than for University of Cambridge International Examinations. The 2008 Report to the Board of Governors calculates that on the operational costs expended on the European Baccalaureate an average per capita cost is in the region of € 655. This is at the high end of what is charged by examination boards in the UK or by the IBO. The very significant difference between the comparisons, however, is that in other organisations the direct costs of examining accounts for only a percentage of total costs with a number of other overhead, staffing, infrastructure, innovation and development costs also accounted for. Marking costs are high – or rather the travel and subsistence costs incurred in the residential marking exercise are high. Somewhat surprisingly, given that teachers contribute draft papers for the European Baccalaureate, the setting costs are also high. 12 subjects incur higher costs than CIE average setting costs at syllabus level. It should also be borne in mind that in the case of the CIE syllabus this typically represents three papers rather than one.

Given that it represents the largest item of spend the current level of marking cost has two major implications for the European Baccalaureate. Examination fees must either be kept at a level that proves inhibitive to the wider adoption of the European Baccalaureate, or development in other areas such as the curriculum and investment in new technologies will be constrained.

Quality

All examinations must be regarded as 'high stakes'. Those examinations which mark the end of secondary education and provide for progression to university are of the highest importance to individuals and impose rigorous standards of accountability on assessment bodies. We consider that the adoption of quality models such as ISO 9001 certification or the quality assurance procedures developed by the Association of Language Testers in Europe (ALTE) would be of value. The European Baccalaureate should also consider establishment of its own Code of Practice to complement the more administratively-oriented focus of the *Arrangements for Implementing the European Baccalaureate.*

The high stakes context of examinations involves a wide range of activities associated with curriculum management, teaching support and assessment delivery. Typically, a cadre of professional and administrative staff are employed to ensure that the annual cycle of activities are managed in full compliance with the tenets of the Code of Practice. Such staff have either professional responsibility for a designated subject or operational responsibility for a specific function. The Evaluation suggests the appointment of a small number of such staff to take forward a number of responsibilities arising from the recommendations of this study.

One other important approach to the improvement of quality within an organisation is through the practice of networking and professional association with other organisations who share a similar mission. Several such organisations exist in the world of educational assessment. Through conferences, network events, journals and training events such associations provide valuable opportunities for the sharing of best practice which we commend to the Board of Governors.

Summary of Recommendations

Section	Recommendation
1.5.5	That the detail of the comparability work be considered on a subject-by-subject basis by European Baccalaureate specification developers. That the common themes of: balance of skills and knowledge; level of material; treatment of material; and progression through topics, all be approached as an across-subject exercise to establish common, though not reductionist, approaches. That Geography is addressed as a subject in particular need of review.
1.6 (a)	To facilitate harmonisation across L2s, it is recommended that a common approach to syllabus design be agreed and implemented across the language sections. A cross-language working party developing this could ensure that best practice from each language section is discussed and a consensus reached that would influence the final design. A shared syllabus design could include a standard structure for all sections of the syllabus and agreed standard introductory paragraphs that are cross-translated. Further, if an outcome focused model, possibly linked to the Common European Framework of Reference for Languages as suggested in 2.4, is developed, the language objectives could be separated by school cycle. This would have the benefit of making explicit the linguistic competence expected in the working language before commencing study of another subject through the medium of L2.
1.6 (b)	Identify common assessment task types that would be effective in testing each skill. Enhance the validity and reliability of the oral examinations by ensuring that the tasks follow a comparable process to the written examinations.
1.6 (c)	The development of common syllabus structures with common assessment objectives and outcomes for all L2s would facilitate the development of common assessment criteria for oral examinations. While the exemplification of performance related to the assessment criteria would clearly differ for each language, a shared understanding of elements to assess could be established by a cross-language group.
1.6 (d)	It is recommended that an agreed standard format be introduced which could specify, for example, the tasks that the student will undertake, the order they should follow, the timings allowed for each section and the interaction of examiners. This could follow the form of an interlocutor frame that guides examiners through the examination and could include standard oral rubrics to structure each part.
1.6 (e)	Cross-language standards setting and subsequent standardisation events involving examiners from all L2 groups could be used to discuss and agree on a common understanding of standards. Such events would require the use of sample oral examinations in languages that are shared by the majority of participants (even if not their L1 or the language they teach) to exemplify how shared assessment criteria (as recommended above) should be applied. The Common European Framework of Reference for Languages can be used to provide a common vocabulary and means of facilitating a common understanding

	of standards expected (note: cross-language standardisation events were introduced as part of the Asset Languages programme).
1.6 (f)	Explore the feasibility of new methods for gathering and retaining evidence generated during oral examinations.
1.7 (a)	It is suggested that this is an area where more empirical work needs to be done within the European Baccalaureate. If the European Baccalaureate has managed to accelerate level 1 and 2 children in the manner suggested by the second PISA report, the specific mechanisms but which this is being achieved are of considerable interest. Likewise, work on the balance of positive and negative impact of 'redoublement' needs to be analysed both in general and in terms of individual students, which is likely to then reveal any specific and unique elements of the European Baccalaureate approach to, and implementation of, 'redoublement' and any enhancement of strategy in this crucial area.
1.7 (b)	We propose that an anchoring test is deployed at the start of the advanced phase and could be used to examine standards across European Baccalaureate schools. We advise that the CEM Centre anchoring test be used (the YELLIS instrument) since this allows linking across to all CEM data and to national qualifications data. Some methodological caveats apply to the assumptions behind the CEM approach, e.g. assumptions that motivational and other factors are constant in relation to continuing attainment. However, despite these limitations, the methodology is considered sufficiently robust for the purpose of linking European Baccalaureate schools and linking into national systems where possible.
2.2	A stakeholder survey should be conducted to identify the scope for the introduction of new subjects into the curriculum. Given the costs of maintaining a wide curriculum in schools, consideration may be given to establishing approved national and international providers of subjects not directly offered in the European Baccalaureate
2.3.2	Consideration may be given to a syllabus review in languages in which a stronger relationship with the CEFR is established. This might be designed with a focus to moving towards the concept of certificating a student's actual level in the CEFR for any of their L2 – L4 options
2.4	Consideration should be given to a curriculum review across all Science syllabuses in which a stronger and more coherent approach to the development of enquiry-based and investigative skill development is established
2.5.3	The introduction of an activity based on the TPE or Extended Essay model might be a valuable and certificated outcome within the European Baccalaureate. Consideration could be given to whether this might itself build on the European identity of the schools themselves.
2.6.2	A stronger relationship between teaching approaches in the European Schools and CLIL initiatives should be established. While this is present in relation to the subjects which are currently examined in a student's L2, the fact that increasing proportions of students are likely to be categorised as 'without a language section' might prompt consideration of this being positioned more positively as 'Content

	and Language Integrated Students'. The fact that the European Schools have established a strong track record in this area might lead to greater involvement in action research initiatives linked to language acquisition in the process of broader curriculum instruction.
2.7	Consideration should be given to the inclusion either of business-related options within the Baccalaureate curriculum or the inclusion of a cross-curriculum approach such as TPE (as proposed in Section 2.7).
3.2.3	We recommend that the European Schools recognise the potential for high validity in the assessment models which the Year 6 and 7 curriculum offers through an integrated delivery of teaching, learning and assessment. The relatively small scale of the current European Schools' operation, together with high levels of teacher experience and of student engagement, makes possible a context in which a high proportion of internal teacher assessment can take place. Teachers should be encouraged to develop teaching and assessment strategies to enrich this context, optimising levels of validity in delivery of the European Baccalaureate.
3.2.4	We recommend that all syllabi are written to contain clearly specified aims and assessment objectives in order to develop a consensus on the knowledge, skills and attitudes which the European Baccalaureate ethos is intended to develop within each subject and that these are clearly articulated in syllabus documents.
3.2.5	That teachers are encouraged to continue to use a broad range of on-course assessments extending their practice into innovative assessments of the type not always possible within large scale national examinations.
	We also recommend that current discourse about whether an assessment is written or oral be replaced by a focus on the purpose of that assessment and the best fit that can be achieved between the form of assessment and its purpose.
3.2.6	The OSGES might give consideration to establishing on-line discussion links between teachers on the Learning Gateway. In particular, these would be designed to explore discussion of teacher-led assessment approaches likely to deliver high levels of validity.
3.2.7	We recommend that marking models, level descriptors or mark schemes, be established for internal and external assessments which link back to clearly expressed assessment objectives for each subject, but which are not necessarily standardised for all subjects in the same way.
3.3.2	We recommend that packs of guidance materials on 'European Baccalaureate Standards' are provided for new teachers. Each should set out clearly the aims and assessment objectives for the subject as well as providing syllabus content, specimen assessment tasks and benchmark scripts at different grade levels together with annotations to indicate why a particular response merited a score of 7.5 for example but not 8.
	Such 'Baccalaureate Standards' induction should then be accompanied by a number of exercises through which a teacher is required to work to ensure that

	they are effectively applying the European Baccalaureate Standard.
3.3.3	Consideration should be given to the way in which new teachers are inducted into the European Baccalaureate Standard and experienced teachers are presented with opportunities to discuss their interpretation of standards with others.
	Opportunities for standards training should be established at the beginning of each year and also before the year 7 Part B examinations.
3.3.4	In respect of the external marking process we recommend (i) that marker standardisation activities should be put in place for all Examiners and that (ii) in addition to the hierarchical processes of quality assurance for marking a process of marking review across subjects be established to provide a mechanism for checking that standards are equivalent across all subjects
3.3.5.3	We recommend the establishment of a minimum entry size in the use of average and distribution statistics in Annual Reporting to avoid the shortcomings of a statistical approach used in the context of unstable small groups. Instead we propose that alternative measures of monitoring marking reliability are established.
3.3.5.4	Consideration might be given to the introduction of a formal cross-moderation research study to review sample scripts and oral tests and thus to make judgement-based analyses for the Board on the reliability of marking in the European Baccalaureate session.
3.4.1	We recommend that the standard analyses prepared for the annual reports be supplemented by a range of further analyses, which should be established as routine protocol.
3.4.3	Consideration may be given to the adoption of alternative approaches to the analyses of exam data to include not only assessment level data (score distributions and summary statistics for the whole cohort), but also subject pairs analysis and item level data.
3.5.3 (a)	Consider calculation of the volume of learning in terms of guided learning hours rather than references to 4 period and 2 period programmes to give end-users and new adopters a clear indication of size.
3.5.3 (b)	Consideration might be given to the notional calculation of subject 'size' to facilitate discussion of its 'currency'. The European Schools might wish to give consideration to a specification of the relationship between curriculum times (for example, 3, 5 and 8 period Mathematics options, the relationship with 4 and 2 periods subject options and subsequent equivalence of standards within the European Baccalaureate.
3.5.4	We recommend that consideration be given to a variation in European Baccalaureate requirements in respect of students who wish to take three Science subjects. It should be possible for them to take the three Sciences as well as mathematics and a higher weighting of the written examinations relative to oral assessment. The European Baccalaureate (Science) might also require that

	students follow practical courses in all three Sciences.
3.5.5	We recommend that the pattern of internal and external assessment be reviewed with a particular emphasis on the internal assessment score of the final written examinations.
3.5.6	We recommend a review of the weighting proposed in the European Baccalaureate Working Group:
	• The proposed reduction of weighting for the Preliminary Mark. • The weighting of written examinations. • The weighting of L1/L2 oral examinations relative to elective subjects. • The combined weighting of L1 written and oral examinations.
3.6.2.2a	If teachers are to remain involved in the question paper setting process, a wide range of materials should be provided to them and sufficient time should be given to enable discussion of student performance in previous sessions, the overall European Baccalaureate and subject goals and assessment objectives, and the Chief Examiners' analysis of the session.
3.6.2.2b	Consideration might be given to the use of a training workshop on item-writing for teachers to ensure that items within a question paper provide appropriate challenge.
	We recommend that the advice to steer the final assessment to year 7 topics be clarified in a way which makes clear that the question paper should sample across the whole programme of study. We also recommend that question styles encourage the synthesis of topics taken from across the two year programme of study.
	If teachers are to continue to be involved in the question paper setting process we recommend that when a syllabus is revised, or a new one introduced, specimen papers are produced, perhaps outsourcing the work to provide assessment exemplars for teaching and learning purposes in advance of the first set of question papers which teachers would be asked to draft.
3.6.2.3	Those responsible for setting papers must increase their use of a wide range of research materials to reflect best practice in their construction of question papers.
3.6.3 a	The European Schools should consider the use of such specification grids in setting assessment tasks that representatively sample syllabus content and assessment objectives and that comply with the assessment structure.
3.6.3 b	We recommend that occasional training be provided for teachers specifically designed to develop the skills of item writing and question paper construction. Their involvement in internal assessment makes this necessary whatever their role in setting questions for European Baccalaureate examinations.
3.6.4	We recommend that teachers continue to be involved in drafting assessments for the final examinations of the European Baccalaureate. However, we propose that

	their role moves to one in which individual teachers are commissioned to produce defined sections of an assessment which sample a syllabus, assessment objectives and scheme of assessment in a manner identified by a specification grid.
3.6.5	We agree with the recommendations of the Working Group that external experts should be contracted to have a role in the question paper setting process. We consider that the specification of their competency should also require, in addition to the features identified by the working group, experience of question paper setting.
3.6.6.1	It may be appropriate to review the practice of providing one written examination in a subject. Two examinations of shorter length deliver both assessment and administrative advantages.
3.6.6.2	The process of the March sign-off of question papers by the Chairman should be accompanied by a review of evidence of a compliance check against syllabus documentation for each subject.
3.6.7	We recommend review of the processes for obtaining copyright for texts and illustrations used within question papers. It may be of interest to review the applicability of materials contained on the CD-ROM of scientific graphics available from Cambridge Assessment.
3.6.8	Consideration might be given to the extended use of secure printer services.
3.7.2	Arrangements should be put in place to ensure that markers are standardised before they begin marking.
3.7.3	A formal analysis of the discrepancy between teachers' and external examiners' marking should be carried out. Formal marking conferences should review discrepancy of this nature.
3.7.4	The process which allows a 20% tolerance between two marks should be reviewed.
3.7.5	We recommend the introduction of a standardisation process to ensure that external examiners across different subjects mark at the appropriate level before they begin their marking in each session.
3.7.6	Consideration might be given to a variation of the current 'double marking' system. Our proposal is that teachers would be asked to draw up a student rank order before sending scripts to Brussels. The rank order would be sent to Brussels sealed. A double external marking would be carried out for any students for whom the rank ordering was interrupted.
	We recommend that all markers for internal and external assessments of the European Baccalaureate undergo a process of 'standardisation'.
3.7.7	We recommend a phased move away from residential marking. From the assessment perspective, there is not strong evidence that the process as currently constituted yields benefits that could not be otherwise replicated.

3.8.3.1	We recommend a review of oral assessment timetabling so that there is the opportunity for a cross-moderation of standards before the issue of results. Our proposals in Section 5.10.5 discuss the way in which this can be done.
3.8.3.2	We recommend a review of the practice by which assessment tasks for oral assessment are designed by the student's teacher. Efficiencies and rigour might be increased by a move toward a centralised design of oral assessments which are provided for teachers in the same way as written examinations.
3.8.3.3	We agree with the reduction in subjects for which there are oral assessments. The removal of the oral in Advanced Mathematics seems sound. However, we consider that the aims and objectives of the syllabuses for History and Geography are so inextricably linked to the ability to communicate that we would recommend retention of oral assessment in these subjects.
3.8.3.4	We recommend review of the duration of oral assessments. Whereas 20 minutes would seem to be appropriate for L1 and L2 assessments for which other performance exists from written papers it would seem an inadequate duration for the assessment of History or Geography.
3.8.3.5	Consideration may be given to the introduction of an oral assessment linked to a presentation which might have been completed as part of a cross-curriculum piece of work (see Section 2.5).
3.9	Consideration may be given to the introduction of an aural assessment developed centrally and provided to ensure that all four language skills are assessed.
3.10	In respect of the roles and responsibilities of stakeholders we recommend:
	• review of teachers' responsibility for the construction of whole question papers • review of teachers' responsibility for first marking (in favour of a rank ordering of candidates) • establishment of a panel of senior examiners or 'jury' able to review and report on the quality of individual examiner, teacher and subject marking, outcomes across all subjects • Inspectors' direct scrutiny of the conduct of examinations in European Schools be delegated to invigilators and observers who are able to report on regulatory compliance to the Board of Inspectors • Inspectors' direct scrutiny of the conduct of marking be delegated to the panel of senior examiners or 'jury'.
4.4.2	That the Guidance Document provided for UK universities by DCSF be swiftly revised. Its information is now out of date. Its reference to equivalent standards in particular in respect of Science A Levels may no longer be appropriate.
4.4.4	That guidance documents for university admissions purposes and European Schools website information are revised using information supplied in the Van Dijk

	study.
4.6	That information on European Schools' websites should be standardised across all schools of the network.
4.7.1	That revised text for the UCAS publication 'International Qualifications' be submitted drawing on information derived from the Van Dijk Report.
4.7.2	We recommend that a dialogue be established with UK NARIC to improve information and understanding.
4.7.3.1	We recommend that an awareness-raising campaign be conducted with key UK university International Admissions Offices.
4.7.3.5	We recommend that European universities are informed of a summary of the Van Dijk outcomes showing high levels of predictive validity.
4.7.3.6	We recommend that a list of officers with assigned responsibility for the European Baccalaureate be compiled for use by European Schools.
4.8.1	We recommend that a dialogue be established with ENIC NARIC to improve information and understanding.
4.8.2.2	We propose that an awareness-raising campaign be conducted with key French institutions.
4.9	Teachers should be advised to check carefully and ensure the applications are sent in good time so that arrangements for late verification can be made where required.
	Given the extent of variation in procedures between and within countries, and indeed in some cases within individual universities, and given the lack of a 'clearing house' for information on this variation, the best resource would be the network of European Baccalaureate schools themselves. The systematic sharing of information regarding procedures in the systems to which their students apply would go a long way towards building up a complete picture for all countries, as a resource on which each school can draw. Such a database would need to be updated on a regular basis.
	OSGES might also produce a briefing document or database to assist teachers in providing guidance to the students.
5.4	Working group proposals for Culham should give consideration to the range of advantages and disadvantages identified in Model A operation to ensure that detailed policy discussion can take place when further expansion opportunities arise.
5.5	Detailed consideration should be given to the range of advantages and disadvantages identified in Model B operation. Further discussion with agencies charged to develop EU multilingualism strategies might be considered.
5.6	Detailed consideration should be given to the range of advantages and disadvantages identified in Model C operation. Further consideration of the model

	presented by the inclusion of the Option Internationale within the French Baccalaureate would be of value – in which a European Option might be provided within the matriculation qualifications of other countries.
5.7	In order to underpin further expansion of the European Baccalaureate, we recommend that a more clearly defined relationship with the Common European Framework of Languages is established.
5.9.3	We suggest that a pilot activity be designed first relating to students' work in class. A study exploring the feasibility of sending work between schools for cross-moderation of marking standards would provide information on students' usage, administrative time involved and the utility of moderating across schools. The study, once evaluated, could then be extended to one subject in Part B examinations, before subsequent ramp-up to the final written examinations.
5.9.4	We suggest that an investigative visit be made to an examination board that employs marking on-line strategies and that the system in use is appraised against requirements for the European Baccalaureate.
5.9.5	Consideration should be given to the development of a project plan by which a move to on-line marking might be effected over a period of three to five years.
5.10.2 (a)	We recommend that the Baccalaureate Unit establishes appropriate benchmarks with other examination boards by which it can appraise whether its costs represent value for money.
5.10.2 (b)	We recommend that the Baccalaureate Unit produces a profile not only of direct costs but also of indirect costs and overheads in order to consider the business model which would fund further expansion of the European Baccalaureate.
5.10.3	We recommend that the Baccalaureate Unit reviews in particular the costs incurred in marking and question paper setting where costs are higher than for the benchmark.
5.10.4	We recommend that a review of the external examiner's actual attendance at all oral examinations be conducted. It incurs high cost. It also introduces significant constraint into the flexibility of the timetable.
5.10.5	We recommend that a review of the March presentation of question papers to the Chairman by the Inspectors be conducted with a view to reducing cost by a different process.
5.10.6	Consideration might be given to standardising the number of meetings required in question paper consideration in order to control costs.
5.10.7	We recommend a review of the practice by which inspectors/subject experts from each country are required to attend question paper meetings.
5.10.8	We recommend a staged move away from residential marking. We propose the establishment of a pilot study to investigate the feasibility of conducting European Baccalaureate marking in examiners' homes.

6.2.1	Consideration may be given to the adoption of ISO as a quality standard.
6.2.3	Consideration may be given to a review of the ALTE standards and to the use of such an approach in the European Baccalaureate.
6.3	The Board of Governors of the European Baccalaureate Examination Board may wish to give consideration to a codification of current, and future, practice with the production of a European Baccalaureate Code of Practice.
6.4	The Board of Governors of the European Baccalaureate Examination Board may wish to give consideration to the appointment of three subject professional officers to take responsibility for a range of assessment-related tasks in the Sciences, Languages, and the Humanities.
6.5	The Board of Governors of the European Baccalaureate Examination Board may wish to give consideration to the establishment of a process of cross-moderation for A & B marks – both by teachers in other schools and by the Lead Examiner.
6.6 :	The Board of Governors of the European Baccalaureate Examination Board might give consideration to the adoption of a process of self-audit to continue to monitor performance and set standards for improvement.
6.7	OSGES may wish to give consideration to membership of European and International Associations of Educational Assessment to create a forum for assessment discussion and the exchange of best practice.
6.8	In parallel with the possible development of bespoke training and of examiners for teachers as item writers and assessors, we recommend a programme of training for all those involved with policy and practice in the European Baccalaureate.

1 External Evaluation of the European Baccalaureate - Comparability analysis

1.1 Subject comparisons – international comparisons on 2008 scripts and specifications

The inception report re-iterated the importance of a study of comparability within the external evaluation:

> *Comparability Study Methodology*
> *A particular focus will be given to a comparability study. As the underlying purpose of the European Baccalaureate is to provide a qualification with portability across national systems and which is of sufficiently high standing to allow effective progression, its comparability with qualifications embedded in other national systems is a crucial issue.*

Cambridge researchers have recently been involved in discussions and developments on state-of-the-art methods in comparability analysis (Newton P, Baird J, Goldstein H, Patrick H & Tymms P (eds), 2007, ***Techniques for monitoring the comparability of examination standards***, Qualifications and Curriculum Authority, London) and selected carefully the methods which would be both possible and appropriate to the external evaluation. The method has necessarily been restricted by the following key factors:

1. the availability of scripts (examination outcomes) for all comparator qualifications
2. the availability of a large number of expert judges
3. the availability of additional statistical information in all settings
4. the distinctiveness of, and differences between, the comparator qualifications.

Factors 1 and 2 precluded the preferred approach of Thurston Paired Comparisons. Given the restrictions, the study focused on judgement by subject experts, their judgements being structured through protocols and reporting instruments which draw from previous studies.

The comparability analysis examined five subjects in depth, rather than more superficial analysis across all subjects taken in the Baccalaureate examination. The subjects analysed were:

1. French L2
2. Mathematics
3. Geography
4. English L1
5. Biology

Experts were contracted in each subject field, both in UK and in other nations where appropriate. Because of the very different qualification reporting arrangements – in terms of the relation between marks and grades – in different qualifications and national systems, the following sampling frame across the European Baccalaureate outcomes was used: 5 scripts in each subject in the top 20%, 5 in the modal point of the mark distribution, 5 above pass and 5 below pass in each qualification.

1.2 Materials available

The comparison was based on some or all of the following, depending on what was available from different national systems: exam papers, mark schemes, scripts, and programme specifications in French, Maths, Geography, English, and Biology. Countries/qualifications included Germany, Ireland, France, England, Sweden, and the International Baccalaureate.

	Exam papers	Mark schemes	Scripts	Programme specifications
European Baccalaureate (International)	**y**	**y**	**y**	**y**
A level (England)	**y**	**y**	**y**	**y**
Advanced Leaving Certificate (Ireland)	**y**	**y**	**y**	**Y**
International Baccalaureate (International)	**y**	**y**		**y**
Abitur (Germany)	**y**	**y**		
Baccalaureate (France)	**y**	**y**	**y**	**y**
National Standards (Sweden)				**y**

Table 1.1

Problems were thus present in discharging the original evaluation specification in respect of:

1. Finland: contact was made with suitable analysts in December 2008 only; a supplementary report will be drawn up, presenting the Finnish findings.

2. Sweden: noting the differences in the assessment model in Sweden (highly devolved assessment model) plus no experts available in Geography.

3. Germany: noting state-based comparisons (Bavaria collaborated with the study, through the helpful offices of Georg Hanf at BIBB) but with no student

scripts for analysis.

4. International Baccalaureate: no student scripts available.

As outlined in the project proposal and inception report, detailed reports have been prepared by expert judges, regarding comparison of the European Baccalaureate outcomes and programme with other national systems. French L2, Maths, Geography, English L1 and Biology were subject to comparison with advanced level programmes, student outcomes and assessment model in respect of: France (Baccalaureate); Germany (Abitur offered in Bavaria); Sweden (Advanced Level National Standards); England (A Level offered by OCR); International Baccalaureate; and Ireland (Advanced Leaving Certificate).

Detailed instructions, analysis instruments and reporting grids were prepared for all experts, alongside packs of student work, and references to all electronic resources needed for the comparisons. Briefings were held face to face or via email and phone to ensure all participants were aware of the detailed requirements of the comparisons (annex).

1.3 Analysis Framework

The analysis showed dimensions of variation in respect of:

1. Subject coverage
   The topics included in the programme requirement, and the models and theories implicitly or explicitly driving the subject content main mode of analysis – comparison of programme specifications.

2. Treatment of subject topics
   Depth of treatment of the subject topics, and the learning activities associated with the blocks of content in the programme and/or the themes running through it
   main mode of analysis – comparison of programme specifications.

3. Expected standards
   The standards which should be demonstrated through the assessment
   main mode of analysis – comparison of programme specifications; assessment specifications and guidelines; comparison of student outcomes.

While European Baccalaureate teachers and the European Baccalaureate office may indeed be interested in the detail of the subject comparisons, the topline analysis of whether the qualification is ‘fit for purpose’ in comparison with other national and international qualifications is the key issue for the European Baccalaureate governance groups. As a result, we have included, as annexes to this report, all of the specific comparative reports. However, this section includes the topline analysis.

This reports the overall ratings derived from the specific comparisons, here arranged into subjects, with attending national and international comparative comment. For example, the Maths comparisons are grouped together and the analysis of the different

national and international qualifications grouped as Maths comparisons. Rather than giving a précis of the detail contained in the annex, this table reports in the following summary form. For each subject, the key issues for the qualification are graded using the following scales:

Specification content

little commonality	coded 3
considerable commonality	coded 2
identical/almost identical	coded 1
cannot determine	coded 0

Differences in treatment

significant differences	coded 3
some difference of no great consequence	coded 2
no differences of consequence	coded 1
cannot determine	coded 0

Variation in standards

serious divergence of standards	coded 3
some divergence in standards	coded 2
identical/can be regarded as the same standard	coded 1
cannot determine	coded 0

Note that the direction of divergence for variation in standards is described in the overview texts and the detailed reports in the annexes.

1.4 Rating Scale

Using these summary dimensions allows European Baccalaureate governance to consider the key issue of the overall international positioning of the European Baccalaureate, thus providing a key to its standing. For more detailed analysis and for those interested in refining the European Baccalaureate specifications, the detailed reports in the annex can be studied.

Subject	Specification content	Differences in treatment	Variation in standards	Notes
French L2				
A level (OCR) England	3	3	1	
Abitur (Bavaria) Germany	3	3	1	
International Baccalaureate	3	3	1	
Advanced Leaving Cert Ireland	2	2	3	
National Standards Sweden	3d	3	0	Stage4 B1CEFR

Subject	Specification content	Differences in treatment	Variation in standards	Notes
Mathematics				
A level (OCR) England	2	3	3	
Baccalaureate France	2	3	1	
Abitur (Bavaria) Germany	2	3	2	
International Baccalaureate	2	2	2	
Advanced Leaving Cert Ireland	3	2	0	
National Standards Sweden	3	2	(1)*0	*at pass level

Subject	Specification content	Differences in treatment	Variation in standards	Notes
Geography				
A level (OCR) England	3	3	3	
Baccalaureate France	3	3	3	
Abitur (Bavaria) Germany	3	3	3	
International Baccalaureate	2	2	3	
Advanced Leaving Cert Ireland	2	2	1	

Subject	Specification content	Differences in treatment	Variation in standards	Notes
English L1				
A level (OCR) England	2	2	2	
International Baccalaureate	3	2	2	
Advanced Leaving Cert Ireland	2	2	2	

Subject	Specification content	Differences in treatment	Variation in standards	Notes
Biology				
A level (OCR) England	1	1	2	
Baccalaureate France	1	1	2	
Abitur (Bavaria) Germany	1	1	3	
International Baccalaureate	1	1	2	
Advanced Leaving Cert, Ireland	2	2	3	
National Standards Sweden	1	1	0	

1.5 Overview of comparisons – with detailed analyses accessible in annexes to this report

1.5.1 French L2

French L2	A level (OCR) England	Michael Featherstone
	International Baccalaureate	Michael Featherstone
	Advanced Leaving Certificate, Ireland	Ms. Catherine Murray - secondary teacher of French and assistant chief examiner with the State Examinations Commission (SEC)
	National Standards Sweden	Gudren Erickson & Kerstin Haggstrom University of Gothenburg Sweden
	Abitur (Bavaria), Germany	Michael Featherstone
	Baccalaureate, France	Not relevant

A level (OCR), England

French L2

There is substantial difference of emphasis between the European Baccalaureate approach to French L2 and the approach in the A level (OCR) in England. This specification meets the nationally-prescribed criteria for the subject and can be considered representative of national examinations in this subject. In contrast to the use of study of literature as a central element, the A level specification has no explicit mention of literary study. Interestingly, despite the very different approaches to the subject, the analyst has stated that '…the standards of language achieved in the sample scripts evaluated are not dissimilar and the level of French produced by the candidates on different courses – even if they are writing in response to different stimuli, on very different subjects and, it would appear, with very different objectives in mind – bears comparison…'.

Abitur, Germany

French L2

While the European Baccalaureate focuses principally on literary-based analysis and learning activity, the Abitur programme requires candidates to cover a wide range of topics of general interest from contemporary French society and culture. This engagement with contemporary culture extends right through the use of stimulus materials and source texts within the Abitur and is in considerable contrast to the European Baccalaureate. A translation exercise is included in the assessment, and the oral assessment has assumed a novel form of a discussion between two candidates. In emphasising functional communication, the Abitur is closer to the qualifications from England, Sweden, and the IB – all of which contrast with the European Baccalaureate. The broad engagement and coverage of the Abitur is in line with approaches in the English, Swedish and IB qualifications, but contrast with the relative narrowness of its range of topics, its source materials and stimuli. It is important to note that the analyst suggests that the level of linguistic competence required to successfully complete the courses compared appears to be broadly similar.

International Baccalaureate (IB)

French L2

There is substantial difference of emphasis between the European Baccalaureate approach to French L2 and the approach in the International Baccalaureate programme and examinations. As with the Swedish national standards and the A level in England, there is a strong emphasis on functional communication in the language and, in comparison to the European Baccalaureate, far less emphasis on the importance of intercultural understanding, and no mention of literary study. The analyst report states: '….the European Baccalaureate programme is … very different from the other two (IB and A level) in the narrowness of its range of assessment tasks, as well as the narrowness of its source materials and stimuli…'. As with A level, the analyst was interested to find that despite the very different purposes and content of the European Baccalaureate, the standards of language attained were broadly comparable.

Advanced Leaving Certificate, Ireland

French L2

The analysis suggests that there is considerable divergence in curriculum purpose, content and treatment. The written papers are different in nature, the European Baccalaureate principally oriented to literature and requires high level analysis and comment of both unseen and set texts at a high level. These elements do not form part of the Leaving Certificate examination. The divergence in standards appears to be acute, with European Baccalaureate candidates requiring (and demonstrating) higher levels of fluency, vocabulary, and competence in idiomatic French. They have considerably advanced ability in respect of setting forward an argument and providing analysis and comment on literary texts.

National Standards, Sweden

French L2

There is substantial difference of emphasis between the European Baccalaureate approach to French L2 and the approach in the Swedish system – note the devolved nature of the curriculum and assessment within the Swedish system, with national standards being the national element of commonality in the system, linking the provision to international standards (level B in the CEFR). While the European Baccalaureate focuses on analysis of literary texts and the exploration of French language through analysis and disputation, the Swedish standards strongly emphasise functional communicative ability such as listening comprehension. The contrast in purpose and specification content renders comparison of standards difficult. The assessment approaches in the Swedish system are distinctly more varied – something shared with the French L2 A level in England and the International Baccalaureate. In common with the analysts of these latter qualifications, the analysts for the Swedish comparison felt that, due to the acute contrasts in purpose and content, it was very difficult to judge the extent to which candidates in the European Baccalaureate would fare in the other systems and vice versa; '…however, an average student at stage 4 would most certainly not be able to reach the pass level. Students at the higher stages of French in the Swedish system, and with grades above a clean pass, would probably be able to pass, provided they had been given instruction within the domain of literary analysis…'.

1.5.2 Mathematics

Mathematics	A level (OCR), England	Sue Croft
	Baccalaureate, France	Dominique Raulin
	Abitur (Bavaria), Germany	Sue Croft
	International Baccalaureate	Sue Croft
	Advanced Leaving Certificate, Ireland	Dr. Aidan Seery, Lecturer in Education and former maths teacher with the International School, Munich
	National Standards, Sweden	Peter Nystrom, UMEA Universitet, Department of Educational Measurement

A level (OCR), England
Mathematics
As a fully modular/unitised programme, A level in England has an elaborate choice structure which makes complex the comparison with the European Baccalaureate. A route with a strong emphasis on mechanics is possible as a route through the A level; this is not possible in the European Baccalaureate. In the probability and statistics options in the A level there is considerable similarity in the level of demand between the European Baccalaureate and the A level. The A level also places a stronger emphasis on algebra, trigonometry, sequences and series, but less focus on functions. As with the IB, the use of a graphical calculator is assumed for all appropriate elements of the programme apart from core maths unit 1, where they are not allowed. Of crucial importance, the overall analysis suggests that the overall level of demand is notably less than the European Baccalaureate; only just matching the year 6 content of the European Baccalaureate.

Baccalaureate, France
Mathematics
There is much stronger emphasis in the French Bac on analysis, justification and proof, and an attendant concern with *communication* using maths. This is reinforced through statements such as 'agility in maths' being a key element of the European Baccalaureate programme, whilst the French Bac programme emphasises analysis and argumentation. However, the analysis suggests that despite the significant differences in overall approach, and some differences in the scope of the programmes (with the European Baccalaureate being broader in topic coverage) the common inclusion of fundamental areas of maths produces examination papers which place a very similar level of demand on students.

Abitur, Germany
Mathematics
Please note that this comparison was limited to an analysis of the set examination paper (Abitur Bavaria), owing to the lack of available student script and the programme specifications. The comparison revealed significant issues regarding: the inclusion of mechanics in the Abitur in contrast to its absence in the European Baccalaureate; the higher level of algebra in the Abitur; and an important difference of treatment regarding higher expectations of interpretation and application of knowledge. This may be mediated, in terms of the level of demand of the qualification, by the facility of German schools to be able to select which questions can be answered. This may contrast with

the European Baccalaureate expectation that students will be well versed in all aspects of the European Baccalaureate curriculum. The lack of strong emphasis on mechanics in the European Baccalaureate is an important note – and sees analysis of A level maths above.

International Baccalaureate (IB)
Mathematics
There is variation in specification and treatment between IB and European Baccalaureate across most elements of the subject. There is a significant contrast in the way in which the IB makes clear 'presumed knowledge' deriving from early phases of study (recognising the different non-advanced programmes which exist in different nations), while the European Baccalaureate makes no explicit reference to this. In respect of topic coverage, there appears greater emphasis in the IB on: trigonometry; matrices (absent from the European Baccalaureate); use of calculus to solve problems; advanced number work in real contexts; and key algebraic topics. The IB gives less emphasis and works at a lower level in respect to vectors, lesser emphasis on probability and statistics, and less emphasis on Functions. This is only ***not*** the case if the IB stats option is taken. A matter of controversy in some countries, the IB places considerable emphasis on the use of graphical calculators as a key aid in solving problems. The greater density of topics coverage suggests that the IB is a more intensive course to teach, with the attendant advantages and disadvantages which this carries: '…the European Baccalaureate programme of study, without a formal recommendation as to teaching time (on topics) would appear to give the teachers more flexibility to plan their own timetable, to teach the course. It may allow the teachers to move at a slower pace than the IB syllabus…'.

Advanced Leaving Certificate, Ireland
Mathematics
In the European Baccalaureate there is greater emphasis on formal analysis and on the theoretical foundations of mathematics, whilst in the Leaving Certificate there is greater emphasis on intuitive approaches and practical techniques. This variation in the statement of orientation and purpose does seem to affect the respective content of the specs. History of maths is an element of the Irish specification and is absent from the European Baccalaureate specification. There are significant differences in specification content, key topics in algebra, geometry and trigonometry are present in the Irish specifications and apparently absent from the European Baccalaureate.

National Standards, Sweden
Mathematics
The pattern of variation between the European Baccalaureate and the Swedish standards is interesting, with inclusions and exclusions in each case. Whilst there is a considerable 'core' of overlap, analytical geometry and probability are not covered in the Swedish standards, whilst the Swedish system is strongly focussed on basic algebra, functions and analysis. Of importance, there are areas of advanced mathematics which are present in the Swedish standards and not present in the European Baccalaureate. The analysis indicates that the Swedish standards stress modelling, problem-solving, reasoning and conceptual understanding to a higher degree than the European Baccalaureate. The analysis has also highlighted the fact that although the specification in the European Baccalaureate emphasises the four

general skills in maths (analysis of problems; manipulation, argumentation, reasoning; communication; generalisation, structuring, synthesising), these are 'not particularly visible' in the European Baccalaureate examination. Critically, the 'pass' standard is considered to be very comparable.

1.5.3 Geography

Geography	A level (OCR), England	Carol Doe
	Advanced Leaving Certificate, Ireland	Ms. Joan Greene - Geography teachers at secondary and assistant chief examiner with the SEC
	Baccalaureate, France	Roger Francoise Gauthier
	Abitur (Bavaria), Germany	Carol Doe
	International Baccalaureate	Carol Doe
	National Standards, Sweden	Not available

A level (OCR), England

Geography

The European Baccalaureate approaches the subject essentially from a regional approach, although themes can be emphasised by teachers. This contrasts with the essentially thematic presentation, delivery and assessment in A level in England. Physical geography has a lower emphasis than human geography in the European Baccalaureate; there is greater balance between these two key bodies of study in the A level. There is much less content in the European Baccalaureate relating to ecosystems, climate and weather, geomorphological processes and landforms. Critically, the European Baccalaureate allows greater option choice in respect of geographical topics than the A level, where the A level limits option choice in order to ensure topic coverage. Skills development is a required, but integrated, element of the qualification whereas in the European Baccalaureate this is a more 'embedded' element. These differences amount to a significant distinction between the European Baccalaureate and the A level. Also a critical issue: '…the European Baccalaureate offers topic choice in Year 6, reducing the actual topics taught, whereas A level and IB only have choice within limits to ensure more balanced coverage of a range of topics…'.

Baccalaureate, France

Geography

It is vital to note that Geography is located within the national curriculum in France in a very different way to its location in the European Baccalaureate. In the French Bac, Geography is bound closely to the teaching of history. Due to its curriculum location and the overall approach to the subject, there are strong contrasts with the European Baccalaureate. It is also important to note that some of the skills which form a key part of the European Baccalaureate are, in the French Bac, taught in the years preceding the Advanced phase. Unlike the European Baccalaureate, which uses Europe as the main context for analysis and learning, with subsidiary global study, the French Bac has principally a global focus, including analysis of development issues. In contrast to the French Bac, the European Baccalaureate adopts a comparatively theoretical stance in relation to key issues and processes. The analysis suggests that the coverage of the curriculum by the European Baccalaureate assessment is more complete than the equivalent assessment in the French Bac, but the curriculum is significantly more

limited, thus possibly resulting in inefficient duplication domain sampling in the European Baccalaureate assessment.

Abitur, Germany
Geography
Please note that this comparison was limited to an analysis of the set examination paper (Abitur Bavaria), owing to the lack of available student scripts and the programme specifications. The analysis suggests that the Abitur requires handling and analysis of more complex sets of data to those encountered in the European Baccalaureate. The degree of interpretation is higher, with greater emphasis on understanding rather than reproduction of knowledge. Increased demand also derives from the greater range of geographical regions to be covered. Combined with a larger range of topics, this allows greater probing, in the Abitur, of different levels of development. These are substantial distinctions in terms of content and demand. In comparison with the European Baccalaureate, the questions in the Abitur examination paper are more sophisticated, are of greater depth and are clearly appropriate to Advanced level study and assessment.

International Baccalaureate (IB)
Geography
As stated above, the European Baccalaureate approaches the subject essentially from a regional approach, although themes can be emphasised by teachers. This contrasts with the essentially thematic presentation, delivery and assessment in the IB. As with the A level comparison, the analysis suggests that the European Baccalaureate is more limited and selective, whilst the IB is broader and deeper in its coverage and assessment of topics. As with A level, there is greater balance between human and physical geography in the IB. Fieldwork is a specified and important element of the qualification. While the IB uses greater physical and environmental underpinning of their human-focussed viewpoints (and thus is aligned with the approach in the A level), the European Baccalaureate takes a largely socio-economic, political view of the subject. As with A level these differences amount to a significant distinction between the European Baccalaureate and the IB.

Advanced Leaving Certificate, Ireland
Geography
The analysis indicates that the European Baccalaureate and the Irish Advanced Leaving Certificate are very close in standard. The impact of the fieldwork component on the overall outcome is significant in the Advanced Leaving Certificate, which highlights the greater emphasis placed on practical work in the Advanced Leaving Certificate in comparison with the ‘preferred’ status of fieldwork in the European Baccalaureate. This is a significant distinction between the two programmes and examinations. There is considerable overlap in topic coverage, but some of this overlap is diminished by the placing of core content of the European Baccalaureate in the elective and optional elements of the Advanced Leaving Certificate. Mirroring to some extent the European focus of the European Baccalaureate, the Advanced Leaving Certificate has an emphasis on analysing Ireland in a European context, but unlike the European Baccalaureate extends the application of key concepts to the global context.

It is important to note that skills-based work is significant in the Advanced Leaving Certificate: '…the (fieldwork) which usually merits a high mark, can bring up a weak script up to quite a high grade…'.

1.5.4 English

English L1	A level (OCR), England	Dr Julian Pattison
	Baccalaureate, France	Not relevant analysis
	Abitur (Bavaria), Germany	Not relevant analysis
	International Baccalaureate	Dr Julian Pattison
	Advanced Leaving Certificate, Ireland	Ms Linda Golden - Secondary teacher - English and assistant chief examiner with the SEC

A level (OCR), England
English L1
The analysis suggests that there is less demand in the European Baccalaureate specification, compared to A level. Elements of the A level such as creative writing and knowledge of the conventions of spoken language are absent from the European Baccalaureate. In the A level, candidates are required to write for a wider range of purposes, drawing on texts and commenting on language features. This includes analysis of the ways in which attitudes and values are conveyed in language and literature. Comparison of genres is an element of the critical approach within the A level, but absent from the European Baccalaureate examination. The notion of text comparison is an important element of the A level, and promoted specifically in the aspects of the specification relating to Shakespeare – this is not an explicit element of the European Baccalaureate.

International Baccalaureate (IB)
English L1
The IB is explicitly a literature, rather than a language *and* literature course. In comparison with the European Baccalaureate, it places a stronger emphasis on literary techniques and strategies. Formal literary criticism is a strong theme of the IB. Non-literary writing is covered in more depth in the IB, and candidates are required to study substantially more texts than the European Baccalaureate. The English L1 for the IB does not in itself require candidates to read widely in languages/literatures other than literature in English – although the nature of the European Baccalaureate programme is explicitly designed to promote this overall. By contrast, IB candidates are required to consider World Literature texts. Overall, the analysis suggests that there is divergence in standard: the IB having a higher level of demand, albeit over a narrower range of texts.

Advanced Leaving Certificate, Ireland
English L1
The Advanced Leaving Certificate in Ireland is carefully tied to expectations and outcomes associated with the preceding stage of education. Based around

comprehending and composing, this integrates the teaching of language and literature. The European Baccalaureate does not include study or assessment of the following elements which are a feature of the Advanced Leaving Certificate: Shakespeare; comparative work of different poets; comparative study of abstract modes; composition in a variety of genres. A long composing assignment is required in the Advanced Leaving Certificate but is not a feature of European Baccalaureate papers. The analysis suggests that the Advanced Leaving Certificate is more detailed in approach and more demanding in respect of analysis.

1.5.5 Biology

Biology	**A level (OCR), England**	Richard Fosbery
	Baccalaureate, France	Dominique Raulin
	Abitur (Bavaria), Germany	Richard Fosbery
	International Baccalaureate	Richard Fosbery
	Advanced Leaving Certificate, Ireland	Mr. Peter Jackson - Secondary teacher of Biology and assistant examiner with the SEC.
	National Standards, Sweden	Gunnel Grelsson, UMEA Universitet, Department of Educational Measurement

A level (OCR), England
Biology
The analysis suggests close alignment between the topics covered in the European Baccalaureate and the A level, albeit with greater sub-topic articulation in the European Baccalaureate. It suggests that there is very close agreement in respect of what is considered to be 'core' to the subject. The options element to the A level does mean potentially that students omit some topics which are required elements of the European Baccalaureate programme. Analysis of the mark schemes suggests that broadly the same the level of knowledge and understanding is demanded, while the analysis of scripts reveals the standards of response on the European Baccalaureate and A level papers are very similar. However, there is an issue in respect of the demanding synoptic assessment which is required in the A level specification, which is not present in the European Baccalaureate scheme.

Baccalaureate, France
Biology
It is important to note that biology is *not* a discrete subject and is studied as part of *scientific formation*. There is considerable divergence in the topics taught in the French Bac and the European Baccalaureate, with the French Bac appearing as a more demanding specification in terms of requirement. Both qualifications require the demonstration of scientific reasoning. In both qualifications the focus appears to be on the construction of scientific analysis rather than experimentation in the subject. The overall judgement is that despite the variation in the form of the specifications, and in the structure of the assessment, the requirements to produce scientific reasoning in the subject mean that the two qualifications are reasonably aligned.

Abitur, Germany

Biology

Please note that this comparison was limited to an analysis of the set examination paper (Abitur Bavaria), owing to the lack of available student script and the programme specifications. The analysis suggests that the Abitur question papers make substantially higher demands of candidates, in comparison with the European Baccalaureate. Similar topics occurred in both qualifications. Although both qualifications test analytical and interpretative skills, there are, however, far more recall-based items in the European Baccalaureate examination. The analyst suggests that the Abitur is considerably more demanding than the European Baccalaureate, and similar in demand to the Advanced Extension Award in England, also placing it as being greater in demand than the A level.

.

International Baccalaureate (IB)

Biology

As with the A level, the analysis suggests close alignment between the topics covered in the European Baccalaureate and the A level, albeit with greater sub-topic articulation in the European Baccalaureate. It also suggests that there is very close agreement in the IB and European Baccalaureate in respect of what is considered to be 'core' to the subject. As with the A level, the options structure to the qualification means that students omit some topics which are required elements of the European Baccalaureate programme.

Advanced Leaving Certificate, Ireland

Biology

The analysis shows considerable overlap between the specifications, with some differences of paradigms – e.g. in analysis of foodchains. Overall, the analysis suggests that there is considerably deeper treatment of topics in the European Baccalaureate in comparison with the Advanced Leaving Certificate in Ireland. The relation between the European Baccalaureate and the comparator here is different to the other comparators, and suggests that the Leaving Certificate is the qualifications which may be poorly aligned in terms of standards and specification.

National Standards, Sweden

Biology

There are some departures between the two specifications – for example, enzymes are not included in biology in Sweden but in chemistry – with some topics present in the Swedish standards and not in the European Baccalaureate specification: the structure and life of viruses; man's relationship to nature from the perspective of the history of ideas, the structure and dynamics of ecosystems; how to determine species, and interaction between Man's organs. One key outcome of the analysis is recognition of the extent to which the openness of the Swedish system is countered by the use of specific text books, which give rise certain topics being 'inevitably covered' even though they are not explicated in depth in the standards. On this, Peter Nystrom from UMEA states: '…The topics in biology commented as "inevitably covered" are topics that are very probable within the specification of the syllabus and/or generally found in textbooks in Sweden. Some of these topics are not expressed explicitly in the syllabus, but in order to do something meaningful with the more generally described goal found

in the syllabus, they more or less need to be covered. Even though there is a certain variety of textbooks in Biology, some of the topics not specified in the syllabus are very likely to be found in almost every textbook. These categories are described by "inevitably covered". There is no mechanism assuring that the topics are covered, but that is the case for all topics because we don't have a central examination and teachers are responsible for grading their students. (It is not clear) how much variation there is in the coverage of different topics in Biology in Sweden. Most likely there is some peer-assessment within the community of Biology teachers putting some pressure on teachers to cover the syllabus, and generally the possibility of some variations in the interpretation of the syllabus is not considered a problem in Sweden…'.

Summary

The detail of the comparisons is included in the annexes to this report, and those interested in the subject-specific issues are referred to those materials. In terms of overall findings, the comparisons reveal a very mixed picture across the different European Baccalaureate subjects – there is no common relationship, such as all subjects being broadly aligned in treatment with other countries, or a consistent relation emerging in terms of demand and standards. None of the analysts identified incoherence or grossly inappropriate content, approaches, or demand in the European Baccalaureate; what emerged were differences linked principally to variations in purpose and commitments. There is one subject, Geography, which appears to require urgent review.

For **European Baccalaureate mathematics**, with some caveats about missing topic areas, the programme and assessment emerged as an appropriate advanced level programme; standards are elevated in comparison with some qualifications.

- For **European Baccalaureate English L1**, the distinctive literary focus of the programme was a key feature: the programme and assessment emerged as an appropriate advanced level programme; there are some issues of lesser demand relating to more restricted curriculum focus.

- For **European Baccalaureate French L2**, the distinctive literary focus marks out the provision and despite the extreme nature of these contrasts, standards of outcome appear comparable, albeit with one exception where the European Baccalaureate clearly exceeds the standards of the other programme, the programme and assessment emerged as an appropriate advanced level programme.

- For **European Baccalaureate Biology**, the significant overlap in content, treatment and demand is only confounded by one instance of a discrepancy in demand which is not in favour of the European Baccalaureate. Broadly, the programme and assessment emerged as an appropriate advanced level programme.

- For **European Baccalaureate Geography**, the positive outcomes of the comparisons of the other subjects in the comparability study were not evident in respect of Geography. A consistent picture of acute discrepancies in content,

treatment and standards points to a need for specific attention to be paid to this European Baccalaureate element.

1.5.5 Recommendation
That the detail of the comparability work be considered on a subject-by-subject basis by European Baccalaureate specification developers. That the common themes of: balance of skills and knowledge; level of material; treatment of material; and progression through topics, all be approached as an across-subject exercise to establish common, though not reductionist, approaches. That Geography is addressed as a subject in particular need of review.

It is vital to note that the comparisons within this study have been made on a subject-by-subject basis. The European Baccalaureate as a whole is a complex curriculum requirement – in a manner similar to the IB and the French Baccalaureate. The subject comparisons are important for progression – e.g. are students prepared for specialist routes in higher education, is the provision outdated in any way etc? However, it is essential to recognise the high level of overall demand of the total programme, particularly deriving from the breadth of the subject requirement, combined with the emphasis on linguistic competence and cultural understanding, with the addition of the multicultural modes of delivery invoked by the composition of the teaching force.

1.6 An important comparability issue: L2 scores across the European Baccalaureate

Significance of L2 learning to the aims of the European Baccalaureate

The development of a second language to a high level of proficiency is a very important part of a student's education in the European Baccalaureate. The annual reports demonstrate that there are consistently differences in average scores across English, French and German for both the L2 and the subject (e.g. History) assessed through the L2 and that these scores vary over time. As stated in the 2007 report 'these variations are hard to explain' and this uncertainty represents a threat to both the validity and reliability of these examinations. This section identifies two key issues as points to address: consistency of standards across L2 subjects and harmonisation of orals for the three vehicular languages.

Validity of L2 assessment

Assessment using L2 as working language: Assessment in L2 can give evidence of language ability and of content-related knowledge and skills. As described in section 3, an assessment is considered valid when it tests what it is intended to test. The development of syllabi as recommended with clear aims and assessment objectives that focus on both the language and subject components for subjects taught through a working language would have a positive impact on the validity of these examinations. The syllabus for each subject should be developed and shared by representatives from each working language to increase the harmonisation across language groups. The most appropriate form of assessment, written or oral, could then be selected for the

most effective evaluation of the objectives which would ensure that the assessment form is ‘fit for purpose’.

Assessment of L2: Language and culture and closely interlinked and authentic cultural diversity forms an integral part of the European School experience. In terms of ensuring comparability of educational outcomes, the objectives of L2 programmes could be brought more closely into line. As the Inception Report noted, “a number of interviewees across the EB stakeholders point anecdotally to the impact of the autonomy exercised within L2 provision in terms of syllabus and assessment methods”. A comparison of the English L2 and French L2 syllabi shows while they follow a similar structure and share some underlying principles including the importance of acquiring all four skills (listening, speaking, reading, writing) and communicative language use, there are a significant number of differences in the way the outcomes for each language are expressed and in the types of activities suggested.

Table 1.2: Comparison of English L2 and French L2 in European Baccalaureate

Section	In common	In English L2 only	In French L2 only	Comments
1.1	general aims			
1.2	subject specific aims intro speech/writing enable learning of human sciences	read accurately for information and enjoyment knowledge and understanding of societies where English spoken	literary and cultural content	
1.3	objectives for reading objectives for writing	objectives for speaking and listening separate	objectives for years 1-3; 4-5; 6-7 objectives for speaking and listening combined	Many differences in the details of the objectives and what they cover
2.1	language functions and purposes for years 1-3, 4-5, 6-7	Listening 1-3: listen/watch for information and pleasure Speaking 6-7: participate in group discussions, role-play and drama	Listening 1-3: distinguish different phonemes in French Reading 6-7: Read and analyse literary works, particularly those on the annual programme	details of objectives vary (some examples given)
2.2		topics		
3.1	section on methods	proposes PPP model	general suggestion to vary activities	
3.2	section on activities	suggests use of variety of activities and lists some learning skills	reading syllabus states number of works to cover each cycle homework syllabus states number and type of written tasks	
3.3		composition of classes	practical activities	
3.4		classrooms		
3.5		resource centres for students		
3.6		teaching resource centre		
3.7		primary and secondary liaison		
3.8		English in the humanities		
4.1	intro to formative and summative assessment		intro to what learners should be expected to be assessed on in global terms	
4.2		marking section – covers similar intro as in French 4.1		
4.2/3	intro to opportunities for assessment	list of opportunities for assessing A and B marks	opportunities for assessing A and B marks broken down by years 1-3; 4-5; 6-7	Some similarities in opportunities for assessment but also some variance
4.3/4	definition of grades – standard EB grid			
appendices		subskills lesson plans	grammar acquisition suggested materials suggested reading (literature)	

1.6 (a) Recommendation

To facilitate harmonisation across L2s, it is recommended that a common approach to syllabus design be agreed and implemented across the language sections. A cross-language working party developing this could ensure that best practice from each language section is discussed and a consensus reached that would influence the final design. A shared syllabus design could include a standard structure for all sections of the syllabus and agreed standard introductory paragraphs that are cross-translated. Further, if an outcome focused model, possibly linked to the Common European Framework of Reference for Languages as suggested in 2.4, is developed, the language objectives could be separated by school cycle. This would have the benefit of making explicit the linguistic competence expected in the working language before commencing study of another subject through the medium of L2.

Reliability of L2 assessment

Task types

Teachers in the European Schools are closely involved in the production of assessment tasks, yet there is evidence that teachers do not have much interaction across language groups (as discussed in section 3). If common objectives and assessment outcomes are shared across languages as recommended, it would also be possible to identify common assessment task types that would be effective in testing each skill. Guidelines for the production of these task types could be produced that identify the knowledge and skills to be assessed and the format of the task. Although these would then be implemented and interpreted on a language-by-language basis, such guidelines would ensure that teachers across language groups were working towards similar parameters and assessing common objectives.

Such a systematic approach to task production is not only relevant to the written tests, but is also applicable to the oral examinations. Currently the responsibility for the writing of oral questions falls to a subject teacher and there is no process of quality approval or moderation of the task (article 6.4.3.2 as quoted in section 3). The validity and reliability of the oral examinations would be enhanced if the tasks used followed a comparable process to the written examinations. Further, where prompts are used to promote discussion, it may be possible to use adaptation of these prompts across languages to further increase harmonisation.

1.6 (b) Recommendation

Identify common assessment task types that would be effective in testing each skill. Enhance the validity and reliability of the oral examinations by ensuring that the tasks follow a comparable process to the written examinations.

Assessment criteria

It is noted in section 3 that only 60% of examiners were reported to use written assessment criteria in the 2008 report of the Survey of Schools and that the recommended target for 2009 is to ensure written criteria are available for all subjects.

1.6 (c) Recommendation

The development of common syllabus structures with common assessment objectives and outcomes for all L2s would facilitate the development of common assessment criteria for oral examinations. While the exemplification of performance related to the assessment criteria would clearly differ for each language, a shared understanding of elements to assess could be established by a cross-language group.

Oral examination procedures

The annual reports highlight the professionalism and positive atmosphere which surround the oral examinations. There is also evidence, however, that the experience for each student within and across schools and language groups is not as uniform as may be desirable to ensure the reliability of the assessment. Variations include the reuse of questions (Report on the 2007 European Baccalaureate) and the participation of the internal and external examiners. The oral examinations require flexibility to allow each candidate to perform to the best of their ability, but this must be balanced with standardisation to ensure reliability.

1.6 (d) Recommendation

It is recommended that an agreed standard format be introduced which could specify, for example, the tasks that the student will undertake, the order they should follow, the timings allowed for each section and the interaction of examiners. This could follow the form of an interlocutor frame that guides examiners through the examination and could include standard oral rubrics to structure each part.

Standardisation

External examiners visiting schools is the method by which there is quality assurance of oral examinations and comparability of standards across schools. It is recommended in Section 3 that a standardisation process for the external examiners should be introduced and this would be desirable for each L2. Similarly, the teacher plays a key role in oral assessments and the Baccalaureate Standards pack suggested in section 3 should also include an element related to the orals.

While these developments would help to ensure standardisation within each L2, additional strategies need to be implemented to increase harmonisation across L2s.

1.6 (e) Recommendation

Cross-language standards setting and subsequent standardisation events involving examiners from all L2 groups could be used to discuss and agree on a common understanding of standards. Such events would require the use of sample oral examinations in languages that are shared by the majority of participants (even if not their L1 or the language they teach) to exemplify how shared assessment criteria (as recommended above) should be applied. The Common European Framework of Reference for Languages can be used to provide a common vocabulary and means of facilitating a common understanding of standards expected (note: cross-language standardisation events were introduced as part of the Asset Languages programme).

It is suggested in Section 5 that the actual attendance of external examiners at oral examinations be reviewed and consideration given to the recording of oral examinations. Not only would this allow for second marking as suggested, but would provide valuable samples that could be used in the above standardisation and orientation activities. Samples used for standardisation purposes could also be retained as evidence for standards over time as described later in this chapter.

An alternative or supplementary approach to obtaining samples for use in the above activities is the video recording of 'oral examinations' carried out for this specific purpose. Production of such recordings does of course rely on an event that is not authentic as it is not conducted for the live assessment of a student's L2 proficiency. There are practical advantages in the quality of the recording as this can be set up to record with high quality visual and sound. Furthermore, examiners can be chosen for the way they conduct the examination, so that the sample not only serves as exemplification of standards but as a model for examiner conduct. High quality recordings can be reused over an extended period of time.

1.6 (f) Recommendation
Explore the feasibility of new methods for gathering and retaining evidence generated during oral examinations.

1.7 Using triangulation data to establish the level of attainment of European Baccalaureate students

As a further part of comparability analysis and in line with the proposal and inception report, the project team has sought to identify data which could be used as triangulation data to establish the level of attainment of European Baccalaureate students relative to those being assessed in different national systems – providing a basis for within-country comparisons (eg European Baccalaureate schools and the national system of the nation in which they are located) and across-country comparisons. Six main studies have been identified as relevant and have been examined in detail: Kelly and Kelly 2006 (The European Baccalaureate – a study of the performance of European Baccalaureate students in Higher Education in the UK and Ireland); the Luxembourg PISA study; the analysis of the academic and professional careers of the European Schools' graduates; the work of UK NARIC; the work of the Eurydice Unit; and PISA 2006 Results for the European School, Luxembourg 1 – meeting in Brussels 14 November 2008.

The project team accessed and reviewed the key analyses which provide a triangulation element:

1. Partial results of the PISA 2006 study at the European School, Luxembourg 1
2. The European Baccalaureate – a study of the performance of European Baccalaureate students in Higher Education in the UK and Ireland, Kelly and Kelly, 2006
3. Analysis of the academic and professional careers of the European Schools' graduates, Policy department B, Structural and cohesion issues, Culture and Education, October 2008
4. The work of UK NARIC

5. The work of the Euridyce Unit
6. PISA 2006 Results for the European School, Luxembourg 1 – meeting in Brussels 14 November 2008.

It was not possible in the time frame of the project to put in place additional assessment instruments as measures which could be used to build an inter-linked, anchored study, of sufficient size and coverage to allow generalisation regarding the European Baccalaureate as a programme. As a result, the team has been able to provide an appraisal and validation of the key existing studies.

1. The PISA study at Luxembourg 1

Due to its size, Luxembourg is over-sampled within PISA. This enables PISA outcomes data to be linked into outcomes data across the European Baccalaureate system. This study indicates that European Baccalaureate students in Luxembourg 1 are highly ranked amongst PISA students:

Reference figures OECD mean all countries = 500 2/3 of students score 440-600					
	European Sch Lux 1	Luxembourg Lycée classique	Best international	Luxembourg (L) Germany (D) France (F)	OECD
Mathematics	L1 574 L2 571	556	550 Hong Kong	L 493 D 503 F 511	500
Science	L1 576 L2 554	552	548 Finland	L 483 D 502 F 511	500
Reading	L1 574 L2 567	549	550 Hong Kong	L 494 D 513 F 519	500
Problem-solving	L1 574 L2 529	555	543 Finland	L 479 D 491 F 496	494

Table 1.3: *Source: Partial results of the Pisa 2006 study at the European School, Luxembourg 1*

The report outlines the relative scores of the different populations, but does not draw specific conclusions regarding the relative educational merit of the European Baccalaureate curriculum. It does make a claim regarding the extent to which the Luxembourg 1 data can be considered typical of the European Baccalaureate system as a whole.

Appraisal and validation

There are a number of methodological issues which require elaboration in order to understand the inferences which can be made from the above data.

Issue 1: the Luxembourg approach within PISA. Luxembourg dramatically over-samples in PISA and we understand from PISA analysts and contractors that Luxembourg links PISA back is able to link back into individual schools.

Anchoring into PISA (the Programme for International Student Assessment (cross-sectional study of 15 year olds)), for the European Baccalaureate system as a whole and then linking at all national systems, is not possible due to (I) the way in which data is anonymised within the OCED data services outputs, with school identifiers removed; (ii) the absence of national examination systems in some nations in which European Baccalaureate schools are located; and (iii) the disjunction of populations and age-groups in the international surveys and the national systems (PISA is a cross-sectional analysis of the attainments of 15 year olds, whereas the age of interest in evaluating the European Baccalaureate system and national systems is 17/18).

Some nations, such as England, do indeed link PISA across to national qualifications via pupil reference numbers, but such comparisons are in the minority in the system. There is also a major methodological question hanging over the use, by some nations (including Luxembourg), of their PISA data in respect of school accountability mechanisms and below-national-level monitoring. The OECD remains concerned about the use of PISA tests and outcomes in this way, since it has the potential to transform PISA processes from low stakes into high stakes arrangements – thus introducing a non-construct related form of variance into the PISA results. This is a serious methodological concern. If there is indeed construct-irrelevant variance – and this has to remain an open question at the present time – then one needs to be very cautious in over-interpreting the exact relation between point score in the Luxembourg schools and those in other nations that do not link PISA in this way.

Issue 2: the cross-sectional and age-specific nature of PISA. PISA surveys 15 year olds. The data cannot in anyway attest to the value added by the post-15 educational provision of the European Baccalaureate. It certainly appears to indicate the distinctive character of the population which enters the post-16 provision of the European Baccalaureate, a population whose ability profile is further concentrated at the upper end by virtue of the ‘holding back’ (‘doubling’) and ‘forced drop out’ strategy of European Baccalaureate arrangements. The fact that the population in the European Baccalaureate school concerned attains in the upper range of the PISA scores at 15 does not constitute evidence that their attainment at 17/18 - at the end of the European Baccalaureate advanced provision - stands in the same relation to the end-achievement of students in other national systems. The evidence in which we would place confidence in respect of this comes from specific comparability work – an approach to which is outlined elsewhere in this report, with the original analysis presented in the annexes to this report – that is, the comparisons of specifications, outcomes, exam papers etc between the European Baccalaureate and A levels, the French Baccalaureate etc. The European Baccalaureate students may indeed be of relatively high general ability (thus scoring well in PISA) but the post-16 provision may still not stretch and challenge students adequately and may still be less than optimum in terms of preparation for Higher Education. Indeed, the

comparability work suggests that there are indeed some important gaps (in Maths) and weaknesses in scope and approach (Geography).

Issue 3: The Luxembourg 1 PISA paper states that '…As the outcomes of the Luxembourg School are virtually undistinguishable from those of the other European Schools as presented in the Annual Report of the Secretary-General, it can be assumed that these results reflect the reality and quality of our school system as a whole…'. It is extremely unlikely that Luxembourg 1 is affected by some constant and systematic bias which means that a high ability profile at Luxembourg 1 is somehow down-rated in the European Baccalaureate system, so that it is constantly and consistently equated with the performance of lower ability and lower attaining students in the rest of the European Baccalaureate system. Thus, the spirit of the final sentence in the quoted paragraph is correct. However, the specific claims of this quote do need to be qualified. As stated immediately above, the PISA assessment of 15 year olds cannot and should not be generalised to the population of 17/18 year olds at the end of their Advanced European Baccalaureate provision; and the Luxembourg PISA data may be contaminated by construct-irrelevant variance. Thus, the implication that the PISA scores show that the end-attainment of the European Baccalaureate '…school system *as a whole*…' (our italics) is not strictly true.

The Luxembourg outcomes need thus to be treated as useful and indicative (of the ability profile of the population entering advanced level provision and of their relative attainment at 15) but not definitive in relation to the comparability or quality of the post-16 provision.

2. Kelly and Kelly 2006, The European Baccalaureate – a study of the performance of European Baccalaureate students in Higher Education in the UK and Ireland

This study shows interesting patterns from the data in the relatively compressed mark distribution within the European Baccalaureate mark range, with clear discrimination in the mark-degree classification relationship, in common with the A level scores (with caveats around the ceiling effects present in A level grading). The claim made by Kelly and Kelly is that European Baccalaureate students gain a higher proportion of high classification degrees, indicating better quality of the European Baccalaureate provision in respect of outcomes and attainment.

Appraisal and validation

This report is indeed illuminating, although some reservation exists regarding the variation in the basis of degree classification within different Higher Education (HE) institutions. The European Baccalaureate population does not distribute into the HE population in the same way as the general HE entry population, and thus different patterns of degree outcome- European Baccalaureate score relation may be a product of this non-matched distribution. This research might be refined by re-analysis of the universities which receive a large European Baccalaureate entry – although this is likely to be adversely affected by (i) the low numbers of European Baccalaureate students in individual institutions; and (ii) restricted range, where those institutions are elite in intake. Using contributing scores as a more sensitive indicator - rather than the coarser measure of degree classification - is problematic due to the very different scoring arrangements used in different institutions; indeed, this often applies to different departments in the same institution.

There is also the issue of 'value added' in respect of the European Baccalaureate. The PISA analysis (see above) suggests that the intake of the post-16 European Baccalaureate provision is high in ability/attainment.

Whilst the results of the comparison of percentages in this report can easily be over-interpreted, they do appear to call into question the basis of the NARIC assumptions regarding equivalence – see below.

3. Analysis of the academic and professional careers of the European Schools' graduates Policy department B Structural and cohesion issues Culture and Education October 2008

This comprehensive follow-up of all European Baccalaureate graduates is a well-grounded empirical analysis of progression. The results point to a high degree of avowed satisfaction amongst European Baccalaureate graduates of their learning experience and attainment in the European Baccalaureate programme; a higher than usual progression to science-related higher education (94%); a lower than usual drop out rate from higher education (c8.8%) and a high rate of participation in higher education in a country with a language other than the native language of the individual student (62%).

Appraisal and validation

The majority of percentage figures used in the analysis in the report are generated from the total dataset of 2,987 European Baccalaureate graduates. The explanatory power of the report is decreased by doing this, since over the period of the operation of the European Baccalaureate system there have been considerable changes in the structure of labour markets, patterns of participation in higher education, and in the performance of national education systems. More suitable treatments of the data would attempt to determine not only overall statistics but trend data. Given that the study possesses more data on more recent graduates from the European Baccalaureate system (with the obvious caveat that stability in careers typically is best measured five years after graduation from higher education) it would be valuable to re-analyse the data using time series data in order to examine progression routes, drop out, etc.

The overall estimated drop out rate from higher education (c8.8%) is significantly lower than the rate in the systems with traditionally low drop out (UK at 22%, source: Audit Commission) and the overall rate in the EU (c30%, source: OECD). This rate associated with European Baccalaureate graduates is in line with the UK system during the 1970s, when only c10% of the 16-19 population progressed to higher education. It is therefore a figure typical of a higher ability entry to higher education.

The report's expression of mild surprise at the strong link between socio-economic status (SES) of European Baccalaureate pupils and their attainment and progression (p51) is itself surprising. SES remains a strong or the strongest predictor of educational attainment in some systems (UK; Germany) (Social Exclusion Unit, Bridging the Gap, 1998, Stationery Office; OECD PISA study, 2002). The important re-analysis which should be undertaken in the data from this study is of the relationship between SES and outcome in the European Baccalaureate system, in the likelihood of drop out (excluding the UK 'independent school' effect). This would establish the power of this relationship in contrast to average figures elsewhere in the EU, and would thus yield a measure specific to the European Baccalaureate system.

This approach also applies to the gendered nature of progression to subjects in higher education and occupational sectors. This is highly gendered in EU economies (Oates T, 2007, Genderwatch) and the issue for the European Baccalaureate is matching the pattern of destination from the European Baccalaureate with routine patterns in EU nations. This requires alignment of the occupational classification for the analysis of the data from this study with the occupational classification used by the OECD, the EU and ILO. Again, this re-analysis would be a valuable exercise to yield a measure specific to the European Baccalaureate.

The report asserts the quality of the European Baccalaureate provision in equipping young people for higher education generally, for stimulating a high rate of progression into science-related higher education, and a high level of international mobility. However, as with the other studies examined in this section, there are limited inferences which can be made regarding the added value of the European Baccalaureate provision and the relative quality of specific subjects. For example, the comparability study undertaken as part of the external evaluation indicates the need for review of the content of Year 6 and Year 7 programmes. The PISA study and the SES data from the progression study provides strong evidence of the elite nature of the European Baccalaureate population. While the progression report provides evidence that European Baccalaureate graduates are progressing well – the challenge is to understand the ***extent*** to which the European Baccalaureate programme is maximising educational attainment for elite groups of students.

As an important footnote, work undertaken by Cambridge Assessment Research and Assessment team has revealed a strong link between attainment in maths and attainment in languages (Bell & Emery 2008). This may be a vital insight into patterns of attainment in the European Baccalaureate, the importance of unintended cognitive linkages in subjects with the programme, and the importance of emphasising and preserving the key emphasis in the European Baccalaureate on extended linguistic competence in foreign languages.

The web-based questionnaire approach in this study appears to have been very successful, and should be periodically undertaken. It is suggested that, with appropriate security and ethical protocols, records be maintained so that further, periodic follow-up studies be a routine part of the European Baccalaureate evaluation processes.

4. NARIC equivalences

Through contact with NARIC staff we understand that detailed studies of equivalence were previously drawn up by NARIC. These are, however, no longer available publicly, and apparently cannot be accessed by NARIC staff themselves. We could not identify or talk to any member of staff who could give any information or insight into the detail of the report. We were thus not able to scrutinise the methodology or findings. However, the NARIC equivalences are treated seriously by stakeholders across the system. Kelly and Kelly include the detail of a response from NARIC on European Baccalaureate equivalences which is critical evidence.

Appraisal and validation

In line with Kelly and Kelly's analysis, we can establish no empirical basis for the NARIC assumption that 'a fail grade at any level in any country will always be comparable to a fail grade at the same qualification level in another country' (letter from Nicholas Everett, Chief

Editor, Information Services UK NARIC to Anthony Swallow, teacher at Brussels II European School, 5 December 2005, quoted in Kelly and Kelly). This would only be true if the populations entered for the qualifications bore strict comparison in terms of attainment profile. There currently is no metric in order to establish a linkage between the populations taking the different qualifications which NARIC links in this statement, which is a key assumption in this statement.

The NARIC assumption is both troubling and damaging. The PISA Luxembourg study, the 2008 Commission study, and Kelly and Kelly suggest that there is adequate empirical evidence which causes serious doubt to be raised regarding the specific assumptions and precise equivalences which NARIC asserts, but neither of these give adequate data for robust re-scaling, in order to replace the NARIC assumption with an empirically-grounded alternative. This would only be yielded by the adoption of a reference/anchoring instrument of the kind outlined below. Such an instrument would in the first administration give a fix on the range of attainment in the European Baccalaureate population, and would take two years to yield the first evidence and linking data regarding the relationship between European Baccalaureate schools and national systems – and in particular, in terms of UK NARIC's key concerns, with A levels within the English system.

This is a very serious matter and the evaluation team believe that NARIC should be explicitly challenged by European Baccalaureate governance to make public the totality of its evidence in support of the assumptions which the organisation has made in respect of its equivalence and tariff statements.

5. Euridyce data on the European Baccalaureate

The Euridyce ranking of schools (van Dijk progession study) in relation to curriculum content fundamentally is based on time devoted to specific subject content. The comparisons are heavily compromised by the different patterns of compulsory schooling across different nations particularly in respect of the relative position of core and optional elements of curriculum frameworks. The European Baccalaureate system emerges as a framework which strongly emphasises second language and science, and places significantly lower emphasis on art-related provision. This is consistent with the conclusions from the findings of the 2008 progression study.

Appraisal and validation

It is unfortunate that the Euridyce data are then transformed into a ranking of schools. This can readily be misinterpreted as a ranking of the ***quality*** of provision, and treated as rankings of the kind presented in PISA, TIMSS and PIRLS. While the Euridyce figures are a very useful comparison of the structural characteristics of different provision, the quality of provision resides in the totality of the curriculum – which includes school effects linked to the nature of pedagogy, school ethos, the impact and form of non-structured elements of the curriculum, etc.

6. PISA 2006 – Results for the European School in Luxembourg 1 – Joint teaching council – meeting in Brussels on 14 November 2008

This report furthers the analysis of PISA data presented in the 'Partial Results' report cited at 1 above. It drills down further into the data and examines:

- differential attainment across the language sections in the main PISA categories - main finding: language 1 students performing significantly higher in maths and sciences than OECD average; language 2, although higher than OECD average, are lower than language 1 students

- gender differences - main finding: that the school shows a pattern of gendered attainment with relatively depressed attainment of female students in maths and science, but again with differences across language sections
- analysis of levels of competency - main finding: that that the EB appears to support high ability students well, that low performance is closely tied to language proficiency in the specific language sections and that the data for levels 1 and 2 suggests that the European Baccalaureate programme works well in 'lifting' students to the required level by the end of the programme

Appraisal and validation

The report correctly identifies the impact of testing 15 year olds in PISA regardless of grade. This does impact on attainment profiles.

The observations in the report regarding explanation for a number of the features of the data appear to rely heavily on close personal knowledge of the students and the programme. The basis for some of the observations appears rather weak, however, and is not in line with research on other programmes and national systems:

On gender differences: the advantage in reading versus science and maths is a common international phenomenon. PISA 2003 and 2006 suggest that the gender gap is widening in the majority of top performing OECD countries, including Finland. The gendered performance in the different subjects should not be explained only in terms of 'motivation', which suggests an over-individualistic cause. The range of causes is complex, and includes cognitive preferences, early and continuing social conditioning, pedagogic styles, assessment models and labour market pressures (Oates T, op cit). Further empirical work (interviews and questionnaires) on the composition of female motivation re science and maths is worth considering.

On maths and science: the report provides further evidence for the strong claim which can be made regarding the high performance of the European Baccalaureate provision in respect of maths and sciences. However, it is very interesting that the results for reading comprehension are lower than for science and maths, given the strong language focus of the European Baccalaureate programme. This could be explained by ceiling effects (or low thresholds) in language required for science and maths learning – i.e. scientific conceptual development can advance at a pace faster than other language-based conceptual development. It is however, an interesting finding, and the issue of breadth of language learning versus depth may well be worth further exploration.

On explanation of key variations: the analysis of 'school environment' (page 4) is misleading. 'School environment' usually refers to individual school ethos and pedagogic models/styles. Here, it really relates to the local composition of education by school type. The first analysis which typically is done in work on student attainment is an examination of the background of students and their prior educational attainment. This tends to provide the major explanation for variation in student attainment.

Finally, a very crucial theme emerges: the issue of repetition. In the US, this is referred to as 'holding back' or 'retention'. The US literature is increasingly focussed on the problems which result from retention in elementary education – with accumulating evidence (from RAND; from the University of Colorado; and from the School of Medicine at University of California) suggesting a strong relationship between high school retention and drop out (a single retention being associated with an 18-28% increase in the chance of dropping out), and apparent increases in graduation rates being frequently associated with encouraging students who are falling behind to transfer to other schools or by placing them in other schools. Analysts in this area suggest strongly the importance not just of examining the data in general, but the impact of retention on specific classes of learners and on individuals. In France, interest in the balance of the negative and positive impact of 'redoublement' is increasing (Pepin B 1998; Goldstein H 2008) with growing concern that the negative impact has not been adequately recognised.

1.7 (a) Recommendation

It is suggested that this is an area where more empirical work needs to be done within the European Baccalaureate. If the European Baccalaureate has managed to accelerate level 1 and 2 children in the manner suggested by the second PISA report, the specific mechanisms but which this is being achieved are of considerable interest. Likewise, work on the balance of positive and negative impact of 'redoublement' needs to be analysed both in general and in terms of individual students, which is likely to then reveal any specific and unique elements of the European Baccalaureate approach to, and implementation of, 'redoublement' and any enhancement of strategy in this crucial area.

Pepin B, 1998, Mobility of mathematics teachers across England, France and Germany: any problems? Paper presented at the European Conference for Educational Research, Ljubljana, Slovenia, 17-20 September 1998.

Goldstein H, 2008, The effects of repetition (redoublement) on the progress of pupils in the first three years of French schooling, University of Bristol.

Using triangulation data to establish the level of attainment of European Baccalaureate students: proposal

The possibility of constructing linked data (across the European Baccalaureate system and into other data on attainment such as PISA) leads us to conclude that further work in this area is worthwhile and technically viable, albeit needing great care in design.

The issue of the relative level of attainment in the European Baccalaureate system is a critical one, and there are two aspects of this issue which the Secretariat might wish to take into account:

- a single empirical study to establish these comparisons
- a continuing facility for monitoring these comparisons.

1.7 (b) Recommendation

We propose that an anchoring test is deployed at the start of the advanced phase and could be used to examine standards across European Baccalaureate schools. We advise that the CEM Centre anchoring test be used (the YELLIS instrument) since this allows linking across to all CEM data and to national qualifications data. Some methodological caveats apply to the assumptions behind the CEM approach, e.g. assumptions that motivational and other factors are constant in relation to continuing attainment. However, despite these limitations, the methodology is considered sufficiently robust for the purpose of linking European Baccalaureate schools and linking into national systems where possible.

The test for 15 year olds would be administered to all European Baccalaureate pupils in the English section. Currently this test is available only in English and a very restricted range of other languages. Development work (critically, standardisation) is underway in respect of different languages and these will come on stream over time. This common instrument would allow all English section students' attainment to be related across all subjects taken, thus providing a linking mechanism across the European Baccalaureate system. This would provide data on commonality of standards. Since this data allows linking into GCSE and A level in the English system, it would also allow linking into the PISA data, as stated above. Where nations link PISA data into national qualifications outcomes, this linking would allow insight into the European Baccalaureate-national system relationship. Administered over time, it would allow monitoring of standards over time, an issue of considerable importance.

Caveats would still exist where the PISA data may be affected by the way in which nations link national and PISA data. No linking would be possible in those systems which do not retain pupil identifiers or do not over sample at a level which allows school level analysis.

Audit Commission (2007). *Staying the course: the retention of students in higher education.* Report by the comptroller and auditor general, HC session, 2006-07, 26 July 200, The Stationery Office.
Bell J. & Emery J. (2008).
Eurydice (2005a). *Key data on education in Europe.* Eurydice Unit.
Eurydice (2005b). *Key data on teaching languages at school in Europe.* Eurydice Unit.
Eurydice (2006a). *Content and language integrated learning at schools in Europe.* Eurydice Unit.
Eurydice (2005a). *Science teaching at schools in Europe.* Eurydice Unit.
Kelly and Kelly (2006). *The European Baccalaureate – a study of the performance of European Baccalaureate students in Higher Education in the UK and Ireland*
Oates T. (2007). *Genderwatch*
Van Dijk (2008). *Analysis of the academic and professional careers of the European Schools' graduates.* Policy Department B: Structural and cohesion policies – culture and education. Directorate General for Internal Policies of the Union.

1.8 Analysis of short and long-term standards

The inception report outlined the purposes of standards comparisons over time in respect of the European Baccalaureate:

> *A particular focus will be given to a comparability study. As the underlying purpose of the European Baccalaureate is to provide a qualification with portability across national systems and which is of sufficiently high standing to allow effective progression, its comparability with qualifications embedded in other national systems is a crucial issue.*

Our recommendation in carrying out the comparability analysis has been to examine five subjects in some depth rather than to extend the analysis across all subjects taken in the European Baccalaureate examination.

This method is contingent on the availability of student scripts. The evaluation team understands that schools retain a number of sample scripts from each year for exemplification purposes, however, following enquiries with schools these were insufficient in number to cover both the mark range and the subjects which are the focus of the standards study, and be adequate for a 'rank ordering' study.

1.8 Recommendation

Therefore, a recommendation of the evaluation is that new script retention protocols be put in place to enable the European Baccalaureate Office to undertake routine enquiries for monitoring standards over time. It is advised that a study of short term standards over time be undertaken every three years. These should be designed as overlapping studies so that they incorporate a means of tracking long term standards over time. However, it is advised that every fourth study (i.e. every twelve years) the study includes a comparison of specifications, papers and mark schemes, and that the rank ordering study for that year is enlarged to include scripts from the provision twelve years previously.

The ideal model for such comparisons is rank ordered comparison across the full mark range and across years – from below the pass mark to the top end of the range. Cambridge Assessment has trialled and investigated the characteristics, manageability and performance of different approaches to such studies, including 'classical' paired comparisons (paired comparisons of scripts), triple comparisons (three scripts) and rank ordering studies (up to ten scripts in each 'comparison pack'). The last method is considered robust and efficient for the study of standards over time in respect of the European Baccalaureate, and it is this approach which is outlined here.

Such studies require repeated judgements by expert judges, who make holistic judgements on the rank order of the allocated scripts. Extraction and judgement of scripts from across the mark range is the only means of establishing exactly what movement in standards is occurring, since movement can be complex combinations of elongation, compression, transposition, etc.

1.9 Outline of rank ordering standards study for the European Baccalaureate

The European Baccalaureate has a distinctive final mark system, with the distribution of scores essentially located above the pass mark of 60.

For a comparison of short term standards in 2011, scripts from the following years would be required as a minimum: 2009, 2010 and 2011. This would represent a study of short-term standards maintenance.

The scripts would be sampled throughout the mark range, from 55 marks to 94 marks. It would be advisable to sample two scripts at each mark point, though only one at each mark point will be needed for the study. The total sample of scripts per study in one subject would thus be 40x3=120. This figure would be multiplied by the different subjects, which are the subject of comparison. Nine scripts (three per year), selected by a set algorithm, would be inserted into packs to enable the comparison, with the judges placing the scripts in rank order. The allocation plan for scripts would avoid packs with opposing extremes of the mark range included, since this is inefficient.

Five judges are considered the minimum for each subject. Fifteen comparisons per script is considered the minimum required for robust overall comparison. Having undertaken previous studies, we can estimate the following rates of comparison to be reasonable within such a design. Scripts would be cleaned of the overall total mark, and as many individual mark totals and examiner annotations as is feasible. They would then be photocopied as many times as necessary for the study (up to five copies of each script might be needed). Assuming each European Baccalaureate script takes approximately five minutes to read, a pack of nine scripts would take 45 minutes. If five judges ranked 10 packs this would be the equivalent of a day's work (though the work could be carried out at home and spread over more than a day to avoid boredom and fatigue). Such studies are frequently rendered complex when qualifications include a variety of components and/or options. This is not a serious problem in respect of the European Baccalaureate, where the examination paper structures are relatively simple in form. For the long-term standards maintenance study, the numerical parameters of the design will need to be adjusted slightly, but if all scripts from each shorter-term study are archived, this will allow great flexibility in the design.

An example plan for allocating scripts to judges is shown in the Excel spreadsheet accompanying this document. The worksheet labelled 'Allocation' shows what scripts each judge would receive in each pack. The worksheet labelled 'Visual' shows how this results in a linking of scripts across the mark range. Each row of this table corresponds to a pack for one judge. The columns of the table are the scripts in the study, in ascending order of mark total. Scripts from year A are shown by a red '1', scripts from year B by a blue '2' and scripts from year C by a green '3'.

A standard protocol should be used for the studies, both in terms of the instructions given to the expert judges and the means by which the data is analysed.

1.10 Data Analysis

The ranked data can be analysed by fitting a latent trait model. To date, a Rasch partial credit model or a Rasch formulation of Thurstone's paired comparison model have been used, as described in Bramley (2005). The result of this analysis is an estimated 'measure' for each script from each year, all on the same latent scale of 'perceived quality'.

The quality of this scale, for example in terms of reliability and fit (of both scripts and judges) should be investigated in the standard way (e.g. references in Smith et al., 2003).

Then the raw mark scale should be related to the measure scale by plotting pairs of (mark, measure) values on a graph. The greater the relationship between mark and measure within each year, the more confident it is possible to be in the validity of the exercise. This is because a poor (or negative) relationship implies that the expert judges were perceiving quality differently from how the mark scheme awarded marks. Different methods of summarising the mark-measure relationship are possible (Bramley et al., 2008), but a simple linear regression of measure on mark has been used in research to date.

From the regression lines (or other best-fit lines) of measure on mark, the mark on test B or C corresponding to a given cut-score (e.g. the pass mark, or 'distinction' mark) on test A can be determined. This mark can be compared with the actual cut-score on test B or C to determine whether and by how much (in terms of raw marks) the performance standard implicit in the test A cut-score differed from that applied to test B or C.

The amount of error[2] in this linking of cut-scores can be assessed by bootstrap resampling of the regression lines, as described in Bramley et al., (2008).

1.11 References

Bramley, T., (2005). A rank-ordering method for equating tests by expert judgment. *Journal of Applied Measurement, 6(2)*, 202-223.

Bramley, T., Gill, T., and Black, B. (2008). *Evaluating the rank-ordering method for standard maintaining*. Paper presented at the International Association for Educational Assessment annual conference, Cambridge, September 2008.

Smith, R.M., Linacre, J.M., & Smith Jr., E.V. (2003). Guidelines for manuscripts. *Journal of Applied Measurement, 4(2)*, 198-204.

[2] Sampling error in the regression line.

2 Curriculum Overview, Years 4 – 7

2.1 Overview

A curriculum may be defined widely in terms of the entire planned learning experience in a school, including everything that promotes learners' intellectual, personal, social and physical development. As such it includes not only a student's classroom experience but also the wider environment in which the values and ethos of the European Schools are developed.

In the light of this definition the most unique, and valuable, element of a student's experience in a European school happens not merely in the classroom but in the multicultural experience that is part of a student's school life. A student's ability to experience their European identity through social and personal relationships, through the interaction with their teachers and peers drawn from the across the European Union, is the defining experience of the educational experience provided. It appears to continue to exercise this influence well into students' academic and professional careers.[3]

The central question of this section of the report is whether there are aspects of the curriculum that could be developed further to ensure that the Schools continue to provide a high-quality educational experience that meets the goals of the European Schools. The section will review:

- Curriculum breadth
- Languages and the relationship with the Common European Framework
- Science in the Curriculum
- Opportunities for cross-curriculum projects, investigations and individual research
- Preparation for the world of work
- The experience of students without a language section and the relationship with the EU Initiative, Content and Language Integrated Learning (CLIL)

2.2 Curriculum breadth

The formal taught curriculum in years 4 – 7 is a sound preparation for further academic study. The proportion of time devoted to compulsory subjects fits well with median figures for other national systems within the EU. Drawing on the information presented in the Van Dijk study, it can be seen that teaching time apportioned to the teaching of maths matches the median EU figure (12%) with a ranking of 13th; and Physics/Chemistry at 13% ranks 8th [4].

Students are presented with a wide range of language choices with L1 – L4 options in the vehicular languages and, at least in the larger schools, many others as well.

---

[3] Graduates attribute the effect of the European Schools curriculum as having a defining effect on their professional career (49% ascribe 'determining impact' and 35% 'some' impact. Analysis of the Academic and Professional Careers of the European Schools' Graduates, October 2008, IP/B/CULT/IC/2007, 073, 13.10.08

[4] Ibid, Table 2.

However, the diversity of language choices and the median ranking of compulsory elements of the curriculum are to an extent offset by a relatively restricted range of subjects that are available. The curriculum of the Irish School Leaving Certificate provides a point of comparison.

The Irish School Leaving Certificate is offered in 34 subjects and is provided at two levels, Higher and Ordinary, in all subjects except in the case of Irish and Mathematics, where an additional (lower) level foundation is provided.

The range of subjects available to students in Ireland but not in the European Baccalaureate enables students to study:

- Business-related courses such as Business and Accounting
- Applied subjects such as Home Economics, Construction Studies, Engineering, Technical Drawing, Agricultural Economics and Agricultural Science
- Arabic, Japanese and Russian as well as European languages.

Choice of Subjects – Irish School Leaving Certificate

Irish	Physics	Mathematics	Home Economics
English	Chemistry	Applied Mathematics	Accounting
Italian	Biology	Art	Business
Spanish	Physics & Chemistry	Music	Economics
French	Agricultural Science	Religious Education	Construction Studies
German	Economic History	Hebrew Studies	Engineering
Russian	Latin	History	Technical Drawing
Japanese	Ancient Greek	Geography	Agricultural Economics
Arabic	Classical Studies		

While the range of subjects made available in the Irish Leaving Certificate offer options not available in the European Schools, it is not itself as wide as in some countries. There are no opportunities to take ICT, Drama, or Media Studies as examination subjects, for example.

An expansion of the Baccalaureate may well require some consideration of a wider range of subjects either offered directly or in partnership with other educational providers.

2.2 Recommendation

A stakeholder survey should be conducted to identify the scope for the introduction of new subjects into the curriculum. Given the costs of maintaining a wide curriculum in schools, consideration may be given to establishing approved national and international providers of subjects not directly offered in the European Baccalaureate.

2.3 Languages and the relationship with the Common European Framework of Reference for Languages

2.3.1 Overview

The development of language proficiency is a very important part of a student's education in the European Baccalaureate. It is important that the standards across different languages are consistent with each other and recognised as having a clear currency for students. In this section, we continue the discussion of issues relating to L2 and suggest that a stronger relationship with the Common European Framework for Languages be established.

2.3.2 Establishing a relationship with the Common European Framework for Languages

Elsewhere in this report, and in the recent report on *Academic and Professional Careers of Graduates of the European Schools* there are proposals that a stronger link might be established with the Common European Framework of Reference for Languages (CEFR). Although some estimation can be made of final levels achieved in the European Baccalaureate, the relationship is currently not strong.

Our mapping indicates that the Baccalaureate standard for L2 is likely to be around C1. However, the establishment of a more formal relationship with the Framework would require the revision of all L2 – L4 syllabuses. The generalised approach of most syllabuses confounds mapping against the CEFR. For example, "Students should be able to listen for a variety of purposes (e.g., for gist, for accurate reproduction) to single or repeated texts, dramatic representations" is the same aim for Years 3 and 5 and it is difficult to map statements of this kind to any particular level. CEFR descriptors provide a framework for greater differentiation by outcome.

The development of a progressive approach in which students could move through levels rather than expressing achievement in terms of L1 – L4 attainment alone might be thought to be of value.

Furthermore, the specificity of definition provided in the CEFR offers a sound framework of learning outcomes which is likely to make delivery into new European Baccalaureate contexts more manageable in a number of linguistic contexts.

Mapping to the CEFR also facilitates comparison with other programmes of study and international language qualifications.

The methodology for mapping to the CEFR could include:

- Using reference level tables from CEFR to develop objectives and assessment criteria
- Comparison by productive skills output (speaking and writing with Council of Europe illustrative samples
- Comparison of tasks set for all skills using criteria from CEFR.
- Content specifications covering key communicative language competence identified in the framework and assessment criteria should be based on criteria statements to facilitate shared understanding of standards.

The importance of languages in the curriculum and increasing target setting for multilingualism across Europe that is likely to be increasingly measured against the CEFR suggests the need for a stronger relationship than hitherto established. This might be thought to be even more the case for L3 and L4 where other documentary evidence in the European Schools curriculum is less well defined than for L2.

.

2.3.2 Recommendation

Consideration may be given to a syllabus review in languages in which a stronger relationship with the CEFR is established. This might be designed with a focus to moving towards the concept of certificating a student's actual level in the CEFR for any of their L2 – L4 options.

2.4 Science in the Curriculum

The expectation that all students in S4 and S5 will follow a strong Science programme establishes good foundations for S6 and S7 study. Its success in terms of influence on a student's subject of study at tertiary level, participation rates on Science-related undergraduate courses and students' success on those programmes points to the positive role it plays in the curriculum.

It is possible that the syllabus documents that provide a framework for such Science education do not fairly represent the positive classroom transactions that take place in Science lessons. It would seem that the process of syllabus review is not well established. Some syllabus content is not as up to date as might be considered necessary. The Physics syllabus is now 12 years old and its assessment overly related to recall objectives. Of more recent date (2005) the Chemistry syllabus contains neither assessment objectives nor a scheme of assessment and only content is specified.

Biology (2002) contains a statement of aims and objectives, but the assessment objectives are not clearly specified. The syllabus refers to questions which must test application of knowledge, analysis and interpretation but there is no clear guidance as to how this might be done. The syllabus does, however specify some learning outcomes.

In many cases the curriculum of years 6 and 7 is less well-defined than in the two preceding years, leading to a position in which teachers begin to teach to the examination rather than syllabus goals - "the curriculum is not the syllabus, it is the examination paper".[5] This fails to realise the full benefit that might be obtained by the Science curriculum.

The report on *Academic and Professional Careers of Graduates of the European Schools*[6] shows a statistical relationship between the importance the European Schools' system attaches to the teaching of Sciences (a solid basis for all, followed later by a specialisation for the most interested student) and the choices made by its students to follow tertiary education in scientific subjects. Insofar as this should be seen as beneficial for the future of society this aspect of progression to higher education is vital. However, the place of Science in the

[5] Comment during interviews with examiners, June 2008, Brussels.

[6] Analysis of the Academic and Professional Careers of the European Schools' Graduates, October 2008, IP/B/CULT/IC/2007, 073, 13.10.08.

curriculum might be interpreted more widely than the university destination it enables, in terms of the enquiry-based and investigative skills which impact upon students' skill development as a whole. In this context the absence of clear syllabus goals concerning specific learning outcomes might be considered an area to be addressed. In particular, it might be observed that opportunities for practical work in the Sciences – an area most designed to build investigative and inquiry based approaches - do not seem soundly integrated with other classroom activities.

2.4 Recommendation

Consideration should be given to a curriculum review across all Science syllabuses in which a stronger and more coherent approach to the development of enquiry-based and investigative skill development is established.

2.5 Cross-curriculum approaches of individual inquiry and investigation

2.5.1 Overview

Many models of cross-curriculum enquiry are offered by Examination Boards to reduce the over-emphasis on a subjects-based curriculum. This issue, in the European Baccalaureate context was referred to by one Chairman who reported that: "There is little epistemological dialogue between the different types of knowledge taught; there is very little effort to deliver a curriculum which is not just the sum of the various subjects taught but which seeks to integrate."[7]

2.5.2 A UK approach – the Extended Project

To address this type of curriculum concern, which is common to many educational contexts, a recent curriculum and assessment project has been developed in the UK drawing on the established experience of the International Baccalaureate assessment of the Extended Essay. The new UK Extended Project requires students to conduct their own research and investigation in a way that crosses subject boundaries and goes beyond the prescribed syllabus. It has a value equivalent to half an A level course – the nearest European Baccalaureate equivalent being equivalence with a two period course of study.

Some recent project titles[8] include:

- Mathematical models used to predict price changes on futures markets
- The toxicity and structure of Prozac
- The clash between Newton's Laws and Einstein's Relativity Theory
- What books make good films?

Titles are chosen by students and work is mentored by tutors over the course of a year.

[7] Report of the Chairman of the 2004 European Baccalaureate Examining Board Ref.: 2004-D-3110 – en-1.

[8] Projects conducted by students from a sixth form college in Hampshire, England, June 2008.

2.5.3 A French model - Travaux Personnels Encadrés (TPE)

Pre-dating this development in England, the French Baccalaureate has also introduced individual project work into its requirements. It has provided an innovation in teaching and learning which encourages breadth and extends students' skills.

The need for such a course came from the university sector concerned with learning outcomes of the laureates on admission to university:

- Students are **not autonomous enough**, nor stimulated to autonomy by the lycée curriculum
- Students do not learn how **to work collaboratively**
- Students do not learn to **extend their studies**, from conception to realisation
- Students **communicative skill** is not developed
- Students **research and investigative skills** are not developed
- Students are not encouraged to think from **an interdisciplinary point of view**.

TPE is concerned with:

Skills development: Autonomy, initiative, project management to achieve product realisation.. In terms of outcomes, students may produce written files, poems, a journal, video, theatre performance, web pages, posters, scientific experiment.

Personal engagement: Students define their topic with the help of their teachers, and decide to work out collectively an individual or collective production, from various documentary resources.

Guided development: Teachers guide the students along the different phases of their research and production, and check the relevance of the selected information referred to the chosen topic.

The TPE is an extended piece of work; there is a personal notebook where a student's learning journey is recorded. The requirement is for an integrative approach across at least two subjects.

The assessment takes into account:

1. the process or learning journey
2. the production itself
3. its presentation.

The assessment comprises:

Internal evaluation: the process is marked 8/20: teachers must evaluate the personal contribution of each student in the case of a collective production. The teachers produce detailed elements of evaluation that will go to the jury, with the proposed mark.

External evaluation: the test is marked 12/20: two teachers (who are not the students' teachers) mark the production itself (accompanied by a personal synthesis that will allow the external examiners to individualize the marking) and the presentation (10 minutes by student,

but with a first time for the group or the individual student to present the realised production, and a second for each student to be evaluated through a talk with the examiners). The proposed mark goes to the jury with detailed rationale.

Test time: in order not to extend the period of examinations, this test is organised to take place before the examination period.

French teachers, who were initially sceptical about TPE, have been progressively convinced of the value of this new approach. Higher education is strongly in favour seeing the development as the introduction of a university approach within the secondary curriculum.

The introduction of elements such as the TPE or the Extended Project offer the European Schools an opportunity to extend the curriculum in a way that is responsive to the constraints of different contexts. A particular consideration in the European Schools context might be to consider whether such a development should build on the European identity of the schools themselves and look at a particular feature of Europe today.

2.5.3 Recommendation

The introduction of an activity based on the TPE or Extended Essay model might be a valuable and certificated outcome within the European Baccalaureate. Consideration could be given to whether this might itself build on the European identity of the schools themselves.

2.6 Students without a language section and Content and Language Integrated Learning (CLIL)

2.6.1 Overview

The number of students without a language section has increased in recent years. This is both because of the wide range of nationalities represented in the European Schools and also because of the decision to close sections with insufficient numbers of students.

The European Schools policy is that students without a language section continue to have access to tuition or support in their mother tongue. This, together with the supportive environment of the European Schools encourages them to perform as well as other students.

2.6.2 The performance of students without a language section (SWAL)

On the evidence of the 2008 Report to the Board of Governors, it is apparent (albeit on a small cohort) that by the time students without a language section (SWAL) take their European Baccalaureate examinations, their average mark in Chemistry, L2, Maths and Physics is better than the average attained by other students. It is necessary to point out the impact of the self-selecting nature of the SWAL students at this stage; however, the outcomes are encouraging.

The performance of SWAL students provides some evidence, perhaps, of the beneficial outcomes of Content and Language Integrated Learning (CLIL) in which the time required for other subjects is safeguarded while maintaining a strong orientation towards language learning in the European curriculum.

Growing awareness of CLIL-related activities, not just in Europe but also globally suggest the opportunity for the European Schools to make a stronger statement of their CLIL policy.

2.6.2 Recommendation

A stronger relationship between teaching approaches in the European Schools and CLIL initiatives should be established. While this is present in relation to the subjects which are currently examined in a student's L2, the fact that increasing proportions of students are likely to be categorised as 'without a language section' might prompt consideration of this being positioned more positively as 'Content and Language Integrated Students'. The fact that the European Schools have established a strong track record in this area might lead to greater involvement in action research initiatives linked to language acquisition in the process of broader curriculum instruction.

2.7 Preparation for the world of work

Today's European Baccalaureate students enter employment at a time when the concept of a 'portfolio' career is key. Flexibility, adaptability, innovativeness and creativity are often regarded as 21st century dispositions in the workplace.

European Baccalaureate students possess a number of vital employability skills. Their multilingualism represents a significant advantage, which evidence shows they are able to use to great effect – almost one in three of European Baccalaureate graduates going on to work in situations where their second language plays an equal role to L1.[9] This together with high standards of literacy and numeracy make a good start in meeting employers' baseline requirements.

However, in terms of other employability skills these must be regarded as implicitly 'caught' rather than explicitly 'taught' in the European Baccalaureate curriculum. The formal and explicit requirements of the European Baccalaureate make no reference to developing, for example, the skills of collaboration, of personal enterprise and critical thinking and problem-solving. Cross-curriculum approaches such as those suggested in Section 2.5 are considered to develop skills not only in preparing students for university but also for the world of work, building a wide range of transferable skills of value in the workplace. In line with our previous recommendation, the Board of Governors may wish to obtain further information on such approaches.

Section 2.2 observed that the curriculum available to European Baccalaureate students is a relatively narrow one. There are no specifically business-related curriculum options, no applied subjects, and no subjects designed to build a 'bridge' between school and work. Responses to the Directors' survey carried out in June 2008 did not, in general, consider that this was a significant omission. One suggested the potential for inclusion within the European Baccalaureate requirements, of participation in a business-related scheme such as Young Enterprise and two considered that a compulsory work experience element might add value to the Baccalaureate as a whole. The majority of responses considered that the level of demand represented by the current qualification was sufficient. Nevertheless, the Board of Governors might wish to consider whether the inclusion either of a Business-related curriculum, or a cross-curriculum approach such as TPE might not achieve the twin goals of preparing students for their university programme and also their working lives.

[9] Analysis of the Academic and Professional Careers of the European Schools' Graduates, Section 4.5.4.

2.7 Recommendation

Consideration should be given to the inclusion either of business-related options within the Baccalaureate curriculum or the inclusion of a cross-curriculum approach such as TPE (as proposed in Section 2.7).

3 An Assessment Review of the European Baccalaureate

3.1 Overview

This chapter considers the European Baccalaureate as an assessment instrument; as the 'tool' used to assess the learning outcomes of the European Schools' curriculum in Years 6 and 7. The key questions addressed by this section relate to the validity, reliability, impact and practicality of the assessment as a whole. This is necessarily, a narrower perspective than one concerned with the quality of the European Schools' mission and curriculum deliverables.

In reviewing the characteristics of the European Baccalaureate from this perspective the fundamental question concerns the extent to which the qualification may be regarded as 'fit for purpose'. There is no one correct model of assessment with which the European Baccalaureate must comply. Our focus is, thus, not to judge the qualification against a 'best practice' model but instead to consider the current context of the European Baccalaureate, to make recommendations for improvement against the dimensions of validity, reliability, impact and practicality and in so doing to make recommendations for those changes which best fit plans for the future expansion of the European Baccalaureate.

The European Baccalaureate model offers a rich and multi-dimensional assessment in which students' achievements can be closely assessed and recorded by their teachers throughout year 6 and the first semester of year 7 using a number of naturally occurring assessment opportunities. In assessment terms this model is most likely to deliver high levels of **validity** and the focus of our evaluation is concerned to optimise this. Our review indicates a number of areas in which the **reliability** of the assessment can be developed further and this, of course, becomes even more important with the proposed expansion of the European Baccalaureate. We use the term **'impact'** to describe the way in which an assessment model produces a 'washback' effect on teaching and learning. While the curriculum of the European Schools avoids many of the negative aspects of 'teaching to the test' which have been seen to characterise some national systems, we make some recommendations for the way in which the positive aspect of the European Baccalaureate's impact on the curriculum can be extended. The **practicality** of the assessment has been a source of concern for many involved with the European Schools and we make a number of recommendations for the way in which the operational effectiveness of examination administration can be improved.

3.2 Validity of the European Baccalaureate

3.2.1 Overview

The relatively small scale of the current European Schools' operation, together with the detailed knowledge which teachers build of their students' achievements and potential, makes possible a context in which a high proportion of internal teacher assessment can take place. Teachers are able to use assessment tasks which are closely related to the processes of teaching and learning and in so doing to tailor their teaching more sensitively to the interests and abilities of different European Baccalaureate students. Classroom teachers are able to use naturally-occurring evidence to build an accurate picture of a student's ability in a subject and to discriminate between the abilities of different students in a teaching group.

Teachers' ongoing observation of students' participation in class and on their written and practical performance properly recorded over the course duration forms a very valuable part of the European Baccalaureate and contributes to the strong validity of the programme. These features contribute to the undoubted strengths of the current qualification.

3.2.2 High levels of teacher experience

Teachers involved in the delivery of the European Baccalaureate tend to be very experienced. Of those working in the European schools for a period of three years or less, only 3% of respondents to a recent survey (conducted as part of the External Evaluation of the European Baccalaureate) had been teaching in their home country for less than two years[10]. A majority (54%) of respondents had had prior experience in a national system of more than ten years. Only two respondents to this survey had never taught in their home country before working in a European School. This level of teaching experience provides a very strong platform of delivery within the current schools and for the evaluation of students' achievement.

The context in which teaching takes place avoids the need to make too strong a distinction between the different dimensions of teaching and learning on the one hand and assessment on the other. This separation can become a characteristic of national examinations, with a number of unintended consequences. Teachers and their students 'conspire' against the examination – teaching to the test and encouraging question guessing strategies to optimise grades. Such behaviours can have the effect of reduced validity. The European Baccalaureate does well to avoid this tendency. In discussions with both students and teachers we observed an admirable absence of the view that assessment was an external hurdle and we were impressed by the sense in which assessment was seen rather as one facet of a coherent school system.

3.2.3 Student engagement

The relationship between learning and assessment is an extremely positive one for students in European Schools. Students see examinations as a natural culmination of their learning programmes. In fact, students interviewed as part of this evaluation expressed some surprise with the approach to examinations taken by the UK-based students at their University (Cambridge), with one student commenting:

"Examination (time) can be quite annoying because students become very pressurised and practice past papers a lot."

The exam culture and the written nature of examinations in the university were in marked contrast to students' previous experience in the European School.

The quality of this educational experience may, perhaps, be effectively summed up by an unsolicited comment received from a teacher during the survey we conducted[11]. Following some (challenging) remarks about the nature of assessment design in the Baccalaureate, the teacher went on to say:

"Having said all this, though, there are two massive advantages to the European Baccalaureate. Every teacher really appreciates:

[10] European Baccalaureate Teachers' Survey, conducted electronically by University of Cambridge International Examinations, October 2008.

[11] Ibid.

- that they are very free to teach what they wish to and in a manner they wish to, within very broad guidelines
- that the students (at the end of the proverbial day) get what they deserve."

These comments, from students and a teacher, point to a strong aspect of the European Schools' education. The quality of experience and of the European Baccalaureate's potential to develop assessment approaches which deliver high validity should be preserved in any future development and expansion plans.

3.2.3 Recommendation

We recommend that the European Schools recognise the potential for high validity in the assessment models which the Year 6 and 7 curriculum offers through an integrated delivery of teaching, learning and assessment. The relatively small scale of the current European Schools' operation, together with high levels of teacher experience and of student engagement, makes possible a context in which a high proportion of internal teacher assessment can take place. Teachers should be encouraged to develop teaching and assessment strategies to enrich this context, optimising levels of validity in delivery of the European Baccalaureate.

3.2.4 Opportunities to deliver greater validity

The validity of an assessment is achieved if it tests the constructs - the knowledge, understanding and skills - which it is intended to test. A valid assessment is essentially concerned with conformity between the purpose and the outcome of an assessment. The development of a rationale - of aims and objectives at syllabus level - is a necessary part of defining purpose in the curriculum. Our review indicated little consistency in this respect. Of the syllabus documents of the European Baccalaureate only a few subjects[12] operate with clearly defined aims and assessment objectives. If a syllabus specifies only the body of content, it is difficult for teachers to develop a consensus on the learning outcomes, the skills and the attitudes which it is intended to develop. The validity of the European Baccalaureate assessment as a whole is accordingly compromised.

3.2.4 Recommendation

We recommend that all syllabi are written to contain clearly specified aims and assessment objectives in order to develop a consensus on the knowledge, skills and attitudes which the European Baccalaureate ethos is intended to develop within each subject and that these are clearly articulated in syllabus documents.

3.2.5 A focus on purpose rather than form

In our review of the policies and procedures relating to the European Baccalaureate, we noted many references to the modalities of assessment – i.e. to whether an assessment was an oral or a written examination and to relative weightings accorded to each in a student's final score. There is a risk that the 'mode' – '5 or 6 written'; '2 or 4 orals' - of assessment

[12] Syllabuses with defined aims and assessment objectives: Economics (published in 2003), Geography (published 2003), Art (2002).

becomes the defining feature rather than its purpose. From the perspective of assessment quality the primary consideration should be concerned more with the 'fitness for purpose' of an assessment than the precise form it takes. Whether an assessment takes the form of a written or an oral examination should be driven by the nature of the content and the assessment objectives. Different types of assessments should be designed to reflect clearly defined assessment objectives in the distinctive subject domains of the curriculum.

A 'written examination' may, in fact, take many forms depending on the nature of assessment objectives; from the 'traditional' three/four hour examination to more innovative forms (portfolios, extended pieces of work and project work) as well. An 'oral' examination may involve presentations, group activities, and demonstrations. A report on the Irish Leaving Certificate[13] expresses this diversity:

"While each subject is assessed by means of written papers, additional assessment techniques are also used in a variety of subjects. These assessments are organised at school level and consist of oral and aural examinations, skills/practical tests, project work, coursework and student tasks; such skill tests may be a live performance which is assessed at the time of performance or the production of an artefact. Projects and coursework are conducted in schools over a period of time under supervision of the teacher and later presented for external assessment. In some cases project work is carried out over a specified period and a report submitted with the written examination".

The wealth of teaching experience possessed by those involved in the delivery of European Baccalaureate syllabuses provides a strong platform for discussion about the design of internal assessments that develop and encourage student learning. The relatively small scale of operation and the sense of community across European Schools make possible assessment practices not practicable on a national scale. We believe there is scope to explore further what types of written or oral examinations might be most appropriate to the curriculum objectives of the European Baccalaureate in order to maximise the validity of teacher-led assessment.

3.2.5 Recommendation

That teachers are encouraged to continue to use a broad range of on-course assessments extending their practice into innovative assessments of the type not always possible within large scale national examinations.

We also recommend that current discourse about whether an assessment is written or oral be replaced by a focus on the purpose of that assessment and the best fit that can be achieved between the form of assessment and its purpose.

3.2.6 Sharing best practice

European Schools offer the potential for a strong and collaborative network in which best practice can be disseminated. A further opportunity to increase the validity of teacher-led assessment could be provided in the use of on-line discussion forums between teachers. The use of the Learning Gateway to promote such on-line discussion between teachers about

[13] Report provided by a member of the Evaluation Team as part of the External Evaluation of the European Baccalaureate, October 2008.

teaching and assessment strategies together with the production of other guidance material for teachers would have undoubted value in this respect.

3.2.6 Recommendation

The OSGES might give consideration to establishing on-line discussion links between teachers on the Learning Gateway. In particular, these would be designed to explore discussion of teacher-led assessment approaches likely to deliver high levels of validity.

3.2.7 Diversity of approaches

In our review of policy documents relating to the European Baccalaureate we have noticed a tendency for definitions of 'good practice' to be associated with establishing a degree of homogeneity across subjects, in which differences between subjects are regarded as needing to be standardised. To take one assessment example, it is not necessary for all subjects to adopt a common approach to the use of marking schemes or level descriptors as recommended by a European Baccalaureate working group. Level descriptors might be regarded as working well in strongly skill-based subjects and mark-schemes in content-based subjects.

A more fruitful approach to the discussion should be concerned with the type of marking model in a subject most likely to deliver the greatest level of validity in an assessment – once more the discussion is about fitness for purpose. For example, if the assessment objectives of the Advanced Mathematics syllabus are concerned with mastery of a body of mathematical content, an assessment mode which requires the ability to articulate clearly one's mathematical thinking in an oral examination might not provide a valid assessment of ability.[14] A marking approach which dealt with skills irrelevant to the construct under investigation is unlikely to be appropriate. The current multi-dimensional nature of assessment of the European Baccalaureate is a very positive feature. We do not believe that different approaches between subjects should be a concern if they are justified by reference to principle. A high-quality curriculum model is confident about differences across subject domains.

3.2.7 Recommendation

We recommend that marking models, level descriptors or mark schemes, be established for internal and external assessments which link back to clearly expressed assessment objectives for each subject, but which are not necessarily standardised for all subjects in the same way.

3.3 Reliability of the European Baccalaureate

3.3.1 Overview

If the validity of the European Baccalaureate can be regarded as a defining strength, this is less securely so in terms of its reliability. The crux of reliability in assessment is, as the Chairman of the 2008 Examination Board emphasises in his report that: "Similar students who

[14] We understand that this is to be removed from options available in oral assessment.

have attended different schools or different language sections must get the same marks and grades. In the same way, similar students who have got their examinations in different years should also receive the same marks".[15]

The number of different judgements made by teachers in assessing student performance has the potential to compromise the reliability of an examination. This is true of all systems involving high levels of internal assessment but it is necessarily an issue of potential concern in the European Baccalaureate. To compound this, the fact that students take their final examinations in different language sections means that a larger numbers of examiners than would otherwise be required assess smaller volumes of student work. Both factors together create a further threat to the reliability of the system.

This section reviews three major dimensions to the issue of reliability in the European Baccalaureate:

- Approaches to the standardisation of teacher assessors
- Marker standardisation and the quality assurance of marking
- Year- on -year reporting of marking outcomes and the monitoring of standards.

3.3.2 Approaches to the standardisation of teacher assessors

As previously mentioned, the key reliability question is whether students of similar ability obtain similar results in different subjects, in different locations and across different examination sessions (that is, that standards over time will not have moved). From the perspective of the European Baccalaureate as an assessment tool, the concern is to establish marker reliability in a context where teachers play a strong role in assessment and where those teachers are themselves drawn from a range of different pedagogic backgrounds.

The fact that teachers within the European Schools are seconded from member states leads to a rich educational environment in which cultural understanding as well as multilingualism is developed. The practice of offering language sections for as many different nationalities as possible provides a secure learning environment for students and safeguards students' return to their own national education systems. However, an unintended consequence of this has, on occasion been referred to as a degree of 'didactic individualism' in teachers' classroom practice.[16]

This is an important observation with respect to examination reliability and becomes even more important in the context of the European Baccalaureate's future expansion. The practice of harmonised examination papers and mark schemes plays an important part in achieving reliability and goes some way to rebutting the claim of 'didactic individualism'. However, a considerable amount of contributory assessment takes place without such structures. A stronger emphasis on ensuring that Year 6 and 7 teachers are inducted into the European Baccalaureate aims, assessment objectives and standards in different subjects would have a positive effect in ensuring a common approach to standards. Training exercises (for example, the blind marking of previously marked scripts) might also be considered.

---

[15] Report of the 2008 European Baccalaureate Examining Board Ref.: 2008-D-2510-en-1

[16] Report of the Chairman of the 2004 European Baccalaureate Examining Board Ref.: 2004–D–3110– en-1
Report of the Chairman of the 2007 European Baccalaureate Examining Board Ref.: 2007–D–2010–en-1

3.3.2 Recommendation

We recommend that packs of guidance materials on 'European Baccalaureate Standards' are provided for new teachers. Each should set out clearly the aims and assessment objectives for the subject as well as providing syllabus content, specimen assessment tasks and benchmark scripts at different grade levels together with annotations to indicate why a particular response merited a score of 7.5 for example but not 8.

Such 'Baccalaureate Standards' induction should then be accompanied by a number of exercises through which a teacher is required to work to ensure that they are effectively applying the European Baccalaureate Standard.

3.3.3 Induction into the European Baccalaureate Standard

The volume of internal assessment carried out by teachers makes it particularly important both that teachers know and apply the appropriate European Baccalaureate Standard, and that the system as a whole has confidence in the use of that standard.

The teachers' survey conducted as part of this evaluation was designed to gain an impression of the way in which they were inducted into teaching at the appropriate European Baccalaureate standard. Responses indicated that while informal opportunities generally existed within a particular language section, they became much less frequent across language sections in the same school and even less so between schools.

- 35% of teachers report that within their language section they frequently discuss and meet with other teachers who are able to answer at first hand questions about school policy and procedure, compared with 7% who report that they rarely have such opportunity (Chart 1).
- Less than 10 % of teachers report frequent opportunities to discuss and meet with teachers across language sections in the school, 10% reporting that such opportunities were rare (Chart 2)
- Only 5% of teachers report that in-service training meetings in their school dealing with professional issues of teaching, learning and assessment occurred frequently, compared with 51% who considered such events rare (Chart 3). Only 1% of teachers reported that such meetings occurred frequently in a school across language sections; 38% of teachers reporting that in-service training opportunities were rare across language sections (Chart 4)
- Professionally-oriented meetings across schools were even less frequent with 65% of teachers reporting them rare within language sections (Chart 5) and 61% reporting them rare across language sections (Chart 6).

The low-frequency of such meetings to develop a common European Baccalaureate Standard in teachers must be seen as a potential threat to the reliability of the qualification. The fact that these teachers are already very experienced in a number of national standards can compound the significance of this issue.

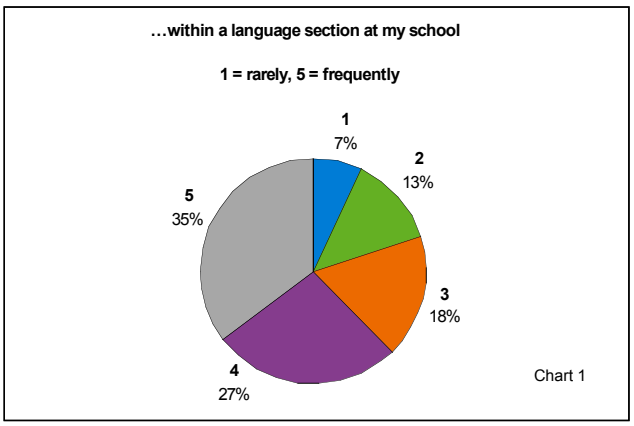
Such initial findings indicate the need for further investigation into the way in which teachers who are given responsibility of assessing students' work to the standards of the European Baccalaureate are inducted into the application of the appropriate standard.

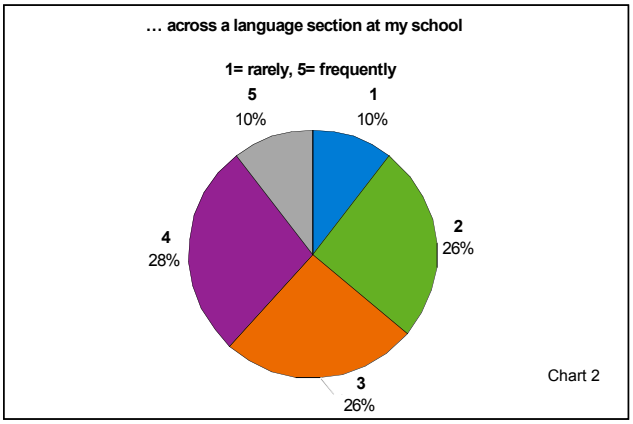
3.3.3 Recommendation

Consideration should be given to the way in which new teachers are inducted into the European Baccalaureate Standard and experienced teachers are presented with opportunities to discuss their interpretation of standards with others.

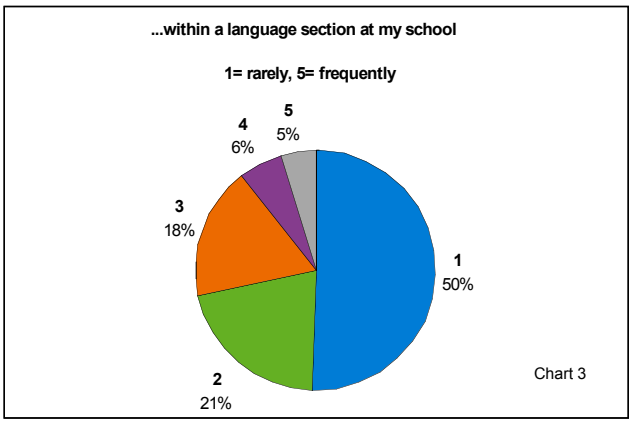
Opportunities for standards training should be established at the beginning of each year and also before the year 7 Part B examinations.

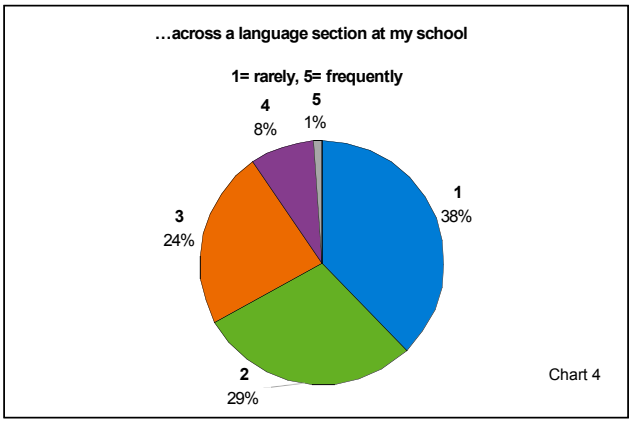
Discussions and meetings with other teachers able to answer at first hand questions about a school's policy and procedures …



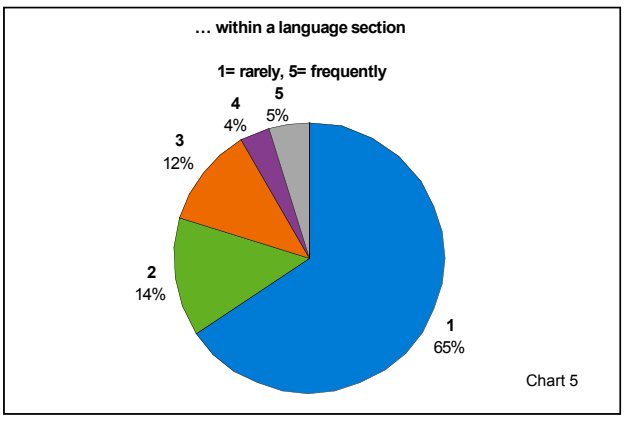


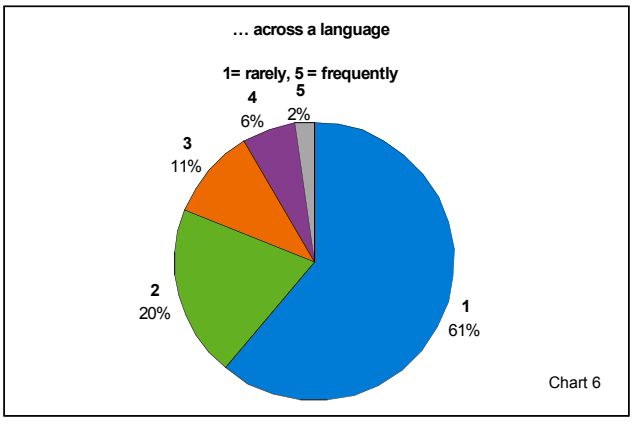
In-service training meetings (dealing with general professional issues of teaching, learning and assessment) for teachers of my subject at my school …





In-service training meetings (dealing with general professional issues of teaching, learning and assessment) with teachers of my subject at other schools …





3.3.4 Marker standardisation and the quality assurance of marking

Examiners selected to take part in the external marking exercise have considerable marking experience. Of those we interviewed during the residential marking exercise in Brussels, most had had extensive marking experience in a national examination system, and some also with the International Baccalaureate. They are most typically recruited by Inspectors and their national referent points are strongly developed. For this reason, we regard it as essential that a standardisation process is developed to guide their European Baccalaureate marking. This might involve an initial trial marking of sample scripts previously marked by the Lead Expert prior to live marking; a formal opportunity for the discussion of standards prior to marking is an essential tool in ensuring reliability.

In terms of the quality assurance of marking across subjects, and year- on -year, we observed no process of marking review. Individual responsibility for ensuring the quality assurance for the marking of each subject lies with an assigned Inspector. But it is not clear precisely where the responsibility lies for ensuring the reliability of outcomes across all subjects in a session. The need for this might be demonstrated by the use of an example drawn from the 2008 session. Result outcomes for this session indicate that average performance declined in all subjects, except one where it remained identical to the previous year (Maths 3). In one subject (Biology) performance increased from an average of 7.31 in 2007 to an average of 7.44 in 2008. This is reported in the November Report.[17] However, a marking review at the end of June would have allowed a consideration of the issues which such a profile throws up. The evidence is that the 2008 cohort does not have higher ability than in 2007 judged on all other subject outcomes. Their performance declines in all subjects other than Maths 3. Nor was there any change in the syllabus which might have explained a different outcome. The hypothesis that would need to be explored is that the assessment tasks for Biology in 2008 have proved easier than anticipated (i.e. they have slipped in standard) or the marking has been more generous than in previous years. A review process before the proclamation of results could improve reliability in this respect.

3.3.4 Recommendation

In respect of the external marking process we recommend (i) that marker standardisation activities should be put in place for all Examiners and that (ii) in addition to the hierarchical processes of quality assurance for marking a process of marking review across subjects be established to provide a mechanism for checking that standards are equivalent across all subjects.

3.3.5 Year-on-year reporting of marking outcomes and the monitoring of standards

While there is no evidence to indicate that the level of reliability in the European Baccalaureate is, in actual fact, less than that for other examinations, the onus of proof lies with the system to demonstrate how reliability is secured. At a structural level this can present a problem for the European Baccalaureate. The total size of the candidate entry

[17] Draft Report on the 2008 European Baccalaureate, Ref 2008-D-2510-en-1, Section 2.7, p47.

(1360)[18] is much smaller than for a national examination and the fact that the entry is split over an increasing number of language sections can make performance trends hard to interpret. The major reporting tool is the Annual Report to the Board of Governors of the European Baccalaureate Examination Board. The 2008 document is considerably more informative than the documents of previous years. It raises reliability questions a number of times: "Are the examination questions becoming easier, year on year? Does the fact that the marks are better mean that the students are better?".[19] It provides careful analysis of the factors that might lead one school's results to be 'better' than another's, with a consideration of the socio-cultural impact which schools will have upon student results: "in conclusion, and although having high marks is an indicator of quality, this cannot mean that one school is better than another."[20] However, the statistical approaches which are used to report outcomes too often frustrate in terms of interpretation: "It should be noted that whenever the number of candidates is very limited the representation in the graphs is not significant."[21] This is indeed the case as:

- cohorts are, for the most part, of a size that do not lend themselves to strong statistical analysis
- variations in the size of entries in the cognate subject makes inter-subject comparisons difficult
- cohort divisions across language sections constrain detailed analysis.

3.3.5.1 Cohort sizes, for the most part, do not lend themselves to strong statistical analysis

In assessment practice, statistical approaches are only considered to be appropriate when applied to a cohort of over a particular size.[22] A cohort of 300 may be regarded as stable in terms of statistical analysis. Even easing this criterion to a cohort size of 200 it is clear that the type of statistical analysis conducted to inform the European Baccalaureate Board of Governors on the quality of a session in terms of its reliability is not always appropriate to the size of the cohort.

Thus, while each of the Annual Reports to the Board of Governors provides a comprehensive analysis of outcomes, illustrated by average and distribution statistics, there are few individual subjects with an entry level which merits statistical analysis. Such statistical techniques, used at national level as a way of ensuring the reliability of examinations, make an implicit assumption that the cohorts of students of one year will be of similar ability and that similar examination outcomes can be taken to indicate that consistent standards have been applied between years and across subjects. The size of the cohort evens out differences between individual students. With a small cohort, however, the distorting effect of individual student ability is such that the initial hypothesis cannot be accepted without the need for further evidence.

---

[18] Draft Report on the 2008 European Baccalaureate for the Teaching Committee (Secondary) meeting ,13 November 2008, Ref 2008-D-2510–dn -1.

[19] Ibid, Section 2.2.3, p22.

[20] Ibid, p19.

[21] Ibid, p28.

[22] We commend the Draft Report on the 2008 European Baccalaureate which acknowledges this point in its narrative – we think for the first time.

3.3.5.2 Strong variations in cognate subject entry volumes make inter-subject comparisons difficult

Statistical approaches are also used within European Baccalaureate Annual Reports to interpret year-on-year trends in performance between subjects in cognate subject areas, irrespective of the number of students entered for each examination. The entry volume for History 2 in the German section, for example, is much lower than in either the French or English sections. The Annual Report for 2005 comments on a change in the average performance of students:

"The two previous years there has been no significant difference, but in 2005 (the standard of) History in German is much lower than French or English".

The report draws the implication that marking is less stable in the German section. Is the volatility of History 2 marking an indication of lack of reliability or is the difference in performance from one year to the next an entirely appropriate outcome given the essential volatility of the relatively small entry for History2 in the German section compared with the French and English sections.

One would, in fact, expect greater instability in standards as the year-on-year impact of differing student characteristics exerts greater influence on the cohort as a whole. It does not, of itself, flag up a problem with the marking in this particular subject. Nor does it, unfortunately, indicate that there is no problem. The crux of the issue is simply that solely on the basis of the statistics one cannot, with confidence, draw this conclusion.

Table 3.1

History 2 Period Entries by Language Section 2003 - 2008

	German	English	French
2008	90	349	200
2007	105	328	163
2006	114	306	209
2005	77	273	170
2004	74	280	172
2003	83	254	1159

3.3.5.3 Cohort subdivisions across language sections constrain detailed analysis

Candidates for the 2007 European Baccalaureate examination were registered within 13 different language sections; as in previous years, performance by subject differs across language sections. And as in previous years the Report comments: "As usual the results differ significantly between language sections".[23] The question is, of course, whether such differences in performance indicate different marking standards and point to unreliability in

[23] For example, see 2005 Report on the European Baccalaureate.

assessment. However, it is also the case that the cohort sizes differ significantly in volume. Enrolment in the English, French and German Language sections fall at one end of the size continuum whereas enrolment to the Danish, Finnish and Swedish language sections fall at the other.

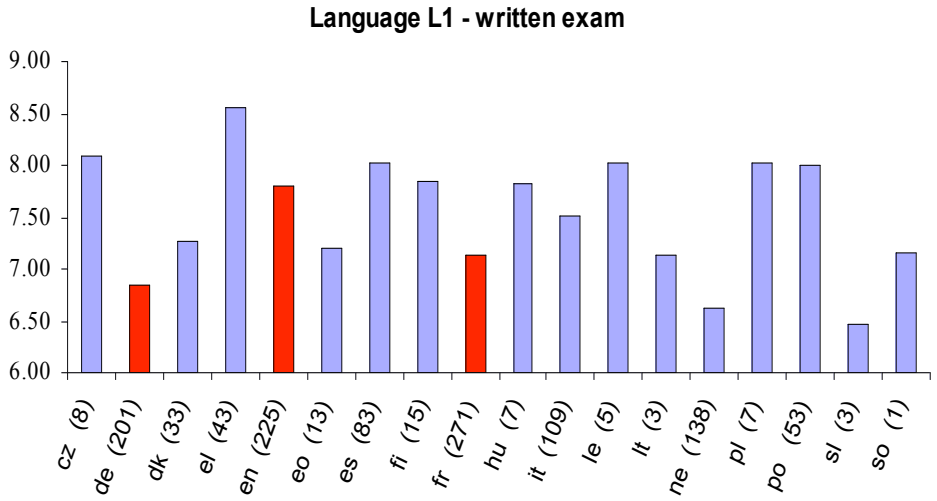


Table: 3.2

One might argue that in the written L1 examination only the entries for German, English and French[24] fall within the scope of appropriate statistical analysis. Given the entry sizes in some other sections, the different outcomes of students in each of the German, English and French sections requires further consideration but statistically little can be said about other language performance. The 'spikiness' of the profile for L1 written examinations in Table 3.1 does not reveal very much about marking reliability at all without reference to further information about the students.

Similarly, while the volume of entries for Mathematics in total falls within the scope of statistical interpretation, the presentation of language section-divided performance re-introduces the problem. Table 3.3 below does indicate the likelihood of weak levels of marking reliability between the English and French sections but in no other sections is it possible to draw this conclusion because of the numbers of students involved.

[24] Report on the 2007 European Baccalaureate, Ref.:2007-D-2110-en-2.

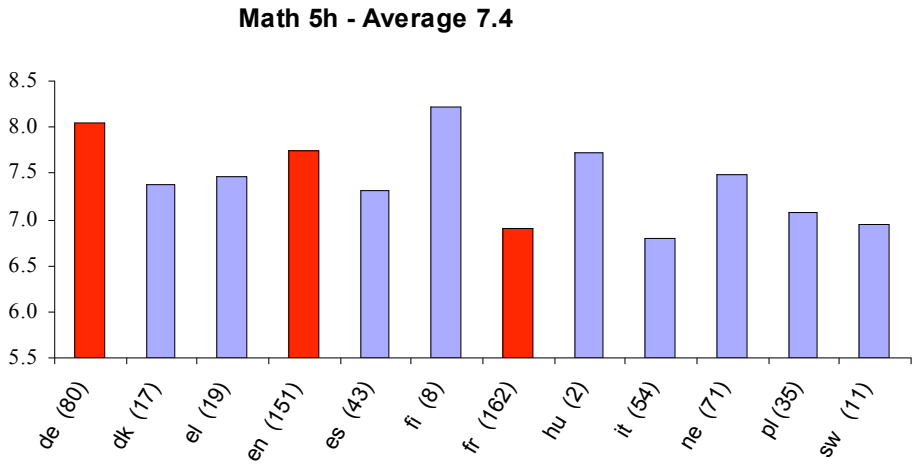


Table: 3.3

The difficulty of drawing conclusions from the statistical profile is not simply because the entry volume covers such a range but also because the spread of schools, and thus of the teachers involved in the marking, is so varied. Whereas there are English, French and German language sections in each of the European Schools only two schools offer Swedish and Danish, and only one offers Finnish. The key concern is not only the statistical one relative to the cohort size but also of the number of 'judges' involved. A replacement of one examiner in a small entry subject can represent a change of 100% of examining personnel in a single year.

The statistics of differing performance in language sections may indeed indicate the need to examine marking reliability further; they do not, in themselves, 'prove' that unequal standards are being applied, for example, across language sections.

3.3.5.3 Recommendation

We recommend the establishment of a minimum entry size in the use of average and distribution statistics in Annual Reporting to avoid the shortcomings of a statistical approach used in the context of unstable small groups. Instead we propose that alternative measures of monitoring marking reliability are established.

3.3.5.4 Alternative measures of monitoring marking reliability

For examinations with low entry volumes where it is difficult to draw conclusions based on average and distribution statistical evidence, other systems of monitoring marking reliability need to be put in place. Such approaches involve cross-moderation processes where the focus is on the collection and use of a number of judgements on scripts across sections, or sessions, or subjects. The key question is: 'Is a 7 on this script equivalent to a 7 on this one?' This might be asked at the beginning of a marking process to standardise markers or as a post-hoc evaluative study. Methodologies in which serial judgements are made about script quality tend to be effective in answering the question about standards. A variant of this approach is to create a rank order via the use of a number of individual judgements about the

better of pairs of scripts to establish agreement with the rank ordering of students which an examination has produced.[25]

The establishment of such formal moderation processes would provide the Board of Governors with confidence that a profile of different results in different subjects or in different schools is attributable to genuine difference in outcomes rather than the lack of marking reliability. An initial research exercise might be used to give a baseline indicator of reliability.

3.3.5.4 Recommendation

Consideration might be given to the introduction of a formal cross-moderation research study to review sample scripts and oral tests and thus to make judgement-based analyses for the Board on the reliability of marking in the European Baccalaureate session.

Although the relatively small sample sizes involved in the European Baccalaureate mean that the statistics cannot be viewed with the same level of confidence as might be achieved in a national-cohort analysis, nevertheless there are some trends which might seem to point to inconsistent standards across different subjects or language sections in the European Baccalaureate examination which cannot be explained simply by the social, cultural and environmental circumstances of the school.

3.4 Suggested analyses of examination data

3.4.1 The European Baccalaureate Office can increase the power of its statistical analysis of the examinations to better monitor the characteristics of the assessment and trends in its characteristics over time.

We have recommended elsewhere (see Chapter 1) the analysis which should be undertaken in respect of comparability between examinations in successive years. We have also recommended that an anchor instrument be used as a means of monitoring changes in the cohort and for examining relative movement in standards in the various components of the examination (see also Chapter 1).

3.4.1 Recommendation

We recommend that the standard analyses prepared for the annual reports be supplemented by a range of further analyses, which should be established as routine protocol.

3.4.2 Assessment level data (as a minimum, total score for each candidate for each assessment)

1. Score distribution and summary statistics (number of candidates, mean and standard deviation of scores), for the whole cohort and broken down by the following categories:

   - Age
   - Gender

[25] See Chapter 1.

- Country/School(with caveats on significance)
- Language of instruction (with caveats on significance)

2. Correlations between scores on each subject of pairs of assessments where there are common candidates.

3.4.3 Item level data (scores on each item for each candidate for each assessment)

1. Omit rate (proportion of candidates not attempting the item).

2. Facility value (the mean mark on the item divided by the maximum mark available).

3. Facility value broken down by quartile (where quartile 1 comprises the lowest scoring 25% of candidates, quartile 2 comprises the next lowest scoring 25% of candidates etc).

4. Facility value might also be broken down by the following categories for those subjects with a significant entry size:

   - Age
   - Gender
   - Country/School(with caveats on significance)
   - Language of instruction

5. Facility value broken down by the following categories within quartile:

   - Age
   - Gender
   - Country/School(with caveats on significance)
   - Language of instruction

6. Correlations between scores on the item and scores on the whole assessment.

3.4.3 Recommendation

Consideration may be given to the adoption of alternative approaches to the analyses of exam data to include not only assessment level data (score distributions and summary statistics for the whole cohort), but also subject pairs analysis and item level data.

3.5 The European Baccalaureate model

3.5.1 Overview

The two preceding sections of this chapter have reviewed the European Baccalaureate from the perspective of its validity and reliability. Well-designed assessments as well as being valid and reliable should also have a positive impact on the curriculum. **Impact** in this context is defined as 'the influence of the test on general educational processes and on the individuals who are affected by the test results.[26] In terms of its impact on the curriculum the structure of the European Baccalaureate model is very well defined to ensure a broad programme of study post 16. The students who achieve the European Baccalaureate undoubtedly demonstrate the effectiveness of the programme.

[26] Studies in Language Testing, Editors: Cyril Weir and Michael Milanovic. Vol 15, 2003, p73.

3.5.2 The European Baccalaureate structure

However, the structure of the Baccalaureate is complex. The options it provides enable students to tailor a programme to their interests and abilities (different length programmes, subject combinations and subject choices and in some cases assessment modes) alongside tightly specified requirements for each of the constituent parts. The complexity of the model can make it difficult to interpret or evaluate against the requirements of other programmes, especially when those other programmes are measured in terms of outcomes such as total learning volume or qualification levels; this becomes important in the higher education recognition process (see Chapter 4)

Universities, employers, parents and other school systems need to understand a qualification in terms of a "currency" that is well-understood. This might relate to:

- **'Size'** measured in terms of the number of subjects studied and the notional learning hours required to attain an appropriate level of performance in a subject
- **Choice** - the extent to which it is regarded as generalist or specialist
- **The weighting of internal/external assessment**.

In each of these dimensions the complexity of the European Baccalaureate model can make it difficult for end-users to interpret its value.

3.5.3 The size of the European Baccalaureate

A qualification can be defined in terms of its 'learning volume'. Within the structure of the European Baccalaureate a unit of learning volume is established through the definition of contact time: each of a student's elective subjects is designed as a four-period a week programme. Contact time refers to the amount of 'classroom time'. In England and Wales an alternative definition of A level volume is used which refers to "the number of guided learning hours" a term used to include both contact time and guided learning time. The volume of a learning programme so calculated is used to determine not only the amount of public funding and the level of resource apportioned to a programme but also as a means of calculating equivalence between the achievements of different students. The number of guided learning hours for an A Level programme is notionally calculated at 360. This becomes the main unit of calculation – a subsidiary course at 50% of an A level is consequently 180 hours. The calculation of contact time (number of teaching weeks × contact hours per week) is straightforward; the extension of this into 'guided learning hours' where students engage in prescribed homework or research activities is less straightforward but can be done. In a context where the opportunities for wider adoption of the European Baccalaureate are under consideration this strategy would make the volume of students' learning programmes clear.

This might be of advantage in the case of the European Baccalaureate, as there is otherwise a risk that students' achievement is undervalued with a consequent negative impact on its currency. For example, the higher education guidance assessment produced by the Department for Children, Schools and Families (DCSF) in the UK,[27] specifically states "Although the European Baccalaureate Science courses have less content/time than GCE A level courses, European Baccalaureate Science students have obtained degree results according to the national distribution".

[27] The European Schools and the European Baccalaureate, Guidance for Universities & Colleges, DCSF.

The document is now quite dated (it was published 2001-02). It is not obvious that the statement on equivalence remains entirely correct. Following the introduction of Curriculum 2000 changes in the UK the syllabus content of European Baccalaureate Science courses appears less but in terms of the volume of time the position is less clear. A student taking 4 period Chemistry, Physics or Biology and taking the Laboratory Option in the subject will be involved in six periods of study a week per subject on the basis of a 31-week teaching year. Contact Time over two years, together with 'directed learning time', i.e. personal investigations and research are, in all likelihood, going to be directly equivalent to the amount of time clearly stated for A level (360 hours to include all guided learning time as well as 'classroom contact time').

This approach of stipulating learning hours, while by no means an exact science, facilitates discussions of equivalence and ensures that full value is given to students' achievement. The Van Dijk study draws attention to the firm grounding which students' receive in their Science education. The Report refers to the high percentage (20%) of European Baccalaureate graduates, about double the EU average of 11% who proceed to study sciences at university and do very well.[28] It is important to the students of the European Baccalaureate as well as its future expansion that this volume of learning is fully recognised.

3.5.3 (a) Recommendation

Consider calculation of the volume of learning in terms of guided learning hours rather than references to 4 period and 2 period programmes to give end-users and new adopters a clear indication of size.

Similar questions about size might be asked about the relationship between Mathematics 5 and 4 period elective subjects and Mathematics 3 and 2 period electives. Whereas universities seem able to understand the overall demand of the Baccalaureate model there is a risk that they might undervalue particular components because of the absence of specificity in these respects. While the main electives and their shorter alternatives maintain a proportionate relationship (4:2 periods) the volume of Mathematics deviates from this: 3 period, 5 period or 8 period options are offered. This gives rise to the question of equivalence. From the point of view of an HE selector the question might be whether a five period a week Mathematics programme:

- offers students a slower route to a level of achievement equivalent to a four period a week programme
- takes students to a higher standard to that reached in their Elective four period subjects.

[28] Analysis of the Academic and Professional Careers of the European Schools' Graduates, October 2008, IP/B/CULT/IC/2007, 073, 13.10.08, p20.

3.5.3 (b) Recommendation

Consideration might be given to the notional calculation of subject 'size' to facilitate discussion of its 'currency'. The European Schools might wish to give consideration to a specification of the relationship between curriculum times (for example, 3, 5 and 8 period Mathematics options, the relationship with 4 and 2 periods subject options and subsequent equivalence of standards within the European Baccalaureate.

3.5.4 Choice

A student's choice, not only of curriculum subjects but also of the assessment mode is a rich source of personalisation in the curriculum. The range of subjects available in European Schools may be considered appropriate to an academically-oriented school system. Students can choose:

- between two and four period options
- between three and five period and four and six period language options
- between 3, 5 and 8 period Maths options
- from a programme of electives and complementary studies.

However, there is a relatively limited choice of subjects. The range tends towards the 'traditional' so there are inevitably some limitations in respect of subjects. From the perspective of higher education requirements this may well be considered appropriate. It is apparent that for some stakeholders the amount of choice within the current system is of significant importance. For example, the restriction of choices concerning whether a subject is assessed by an oral or a written examination at the end of Year 7, which forms part of the European Baccalaureate Working Group's proposal, has caused concern to Interparents:

"Pupils would then only be able to choose from a limited range of subjects, prejudicing the existing free option choices within the system, and probably leading to fixed packages of pre-combined subjects".[29]

While we understand the desire to promote a student's breadth of choice, we agree with the proposal that the choice about whether to be assessed in a subject by an oral or by a written examination is not a feature that adds value to the European Baccalaureate as an assessment tool or to the recognition levels awarded by higher education institutions.

The ratio of core: elective subjects, however, can constrain student choice. This is particularly so in the case of science students for whom the pressures on curriculum time appear most acute. In 2008 20 students needed to sit an additional written paper, mostly in order to offer the three sciences and mathematics for admission to Medical School. It is important that provision be made for calculation of European Baccalaureate performance which takes account of this. However, for the overwhelming majority of students the current requirement of five written examinations appropriately meets their needs. One option which might be considered is to develop a specialised option of the European Baccalaureate which offers a science orientation, in which the expectation would be that six written examinations,

[29] Interparents Memorandum, 2008, p2.

together with a complementary programme of all three Practical Science subjects would be taken. In order to effectively reflect the volume of achievement which this represents an alternative European Baccalaureate weighting might be considered.

3.5.4 Recommendation

We recommend that consideration be given to a variation in European Baccalaureate requirements in respect of students who wish to take three Science subjects. It should be possible for them to take the three Sciences as well as mathematics and a higher weighting of the written examinations relative to oral assessment. The European Baccalaureate (Science) might also require that students follow practical courses in all three Sciences.

3.5.5 The use of internal assessment and weighting of internal and external assessment in the European Baccalaureate

Assessment by teachers in Year 6 and the first semester of Year 7 makes an important contribution to the quality of education in the European Baccalaureate. However, the way in which it contributes to a student's final score is not always well understood. There is a degree of complexity in the European Baccalaureate model with respect to its assessment modes. A common typology used in assessment systems is to distinguish internal/external assessment, final (summative) and continuous assessment. Of the two elements referred to in descriptions of the European Baccalaureate, the first contains the second and the second contains the first. The calculation of the final assessment is not easily understood and may constrain the future expansion of the European Baccalaureate.

3.5.5 Recommendation

We recommend that the pattern of internal and external assessment be reviewed with a particular emphasis on the internal assessment score of the final written examinations.

3.5.6 Current proposals for reform of weightings

The table below sets out the proposal for reform. In this, it is proposed that 25% of marks are awarded for preliminary assessment and 75% for final assessment. The preliminary assessment represents an internal assessment whereas the final assessment is comprised of internal and external assessment in a way not always well understood by those stakeholders with less than a detailed knowledge of the European Baccalaureate. In fact, the final assessment is sometimes described without reference to its internal assessment.

The European Schools' website, for example, simply refers to the 'double marking' of the assessment – "Close scrutiny of the Examining Board, which demands double correction and may require a third, guarantees the high level and quality of the Baccalaureate" - without making clear that the double marking also represents a further internal assessment input into the final examinations.[30]

[30] www.eursc.eu/index.php?id=138

	Assessment type	Weighting	% of the final mark
A mark	internal	100%	10%
Written part examinations (B mark)	internal	100%	15% (preliminary mark= 25%)
6 written examinations	internal + external	1/3 2/3	50%
2 oral examinations	internal + external	1/3 2/3	25% (75%)
Proportion of internal assessment Proportion of external Assessment			50% 50% (100%)

Our observations on the recommendations of the European Baccalaureate's Working Group are as follows:

- The value contributed by the internal assessment during Year 6 and the first semester of year 7 is a defining characteristic of the model which produces high levels of validity. The current weighting is 40% and the proposal is to reduce this to 25%. Given the benefits in terms of positive impact on the educational experience and the potential validity provided by teacher-led assessments specified by the European Baccalaureate Preliminary Mark, the weighting of 40% might not be considered inappropriate. The proposal to reduce the weighting to 25% may possibly undervalue the contribution made by this component of the European Baccalaureate.

  Bearing in mind the 'impact' on the system as a whole of decisions about weighting, a reduction in the value of internal assessment could have serious consequences. Students may decide to give less emphasis to their first year examinations, for example, as it would contribute a relatively low proportion of marks.

- Students' final examinations, written and oral, represent an important part of their European Baccalaureate. Taking all final examinations together, a weighting of 60% relative to 40% for the preliminary mark would seem appropriate.

- The proposal suggests a 35%:24% ratio of six subjects to the two language orals. For students who are preparing for Science, Medicine and Engineering courses the

proposal would entail less weighting for a Science subject than for the oral component alone of one of the language assessments; whereas the Science subject would contribute a major component of their pre-university study. Under the proposal language assessments would in total contribute approximately 42% of the proposed 75% weighting of marks for the final examination. We consider that the weighting of the written examinations might, in this respect, be understated.

- It might be of interest to consider that in University of Cambridge International Examinations oral marks are reported separately but integrated with the written examination mark in the final grade achieved in language examinations.

  The heavy weighting for oral assessment in L1 and L2 is noted by parents: "It seems unreasonable to us to make it possible for students to take examinations only in languages but at the same time to calculate them as being worth 25% of the final mark. We find the respective weight of languages – natural sciences – social sciences unacceptable as it gives students who are scientifically gifted far worse chances".[31]

- The proposal for weighting the outcomes of L1 and L2 orals at 25% of the total European Baccalaureate may be considered to distort the breadth of student's learning and in turn to reduce the extent to which the European Baccalaureate as a whole is well understood.

- The volume of the contribution of L1 assessment is significant, in the proposed new structure, when the written and the oral examination are taken together. Mother tongue assessment is undoubtedly important; however, the heavy weighting of both written and oral assessment may be considered to distort its place in the curriculum and create an inequity between those students whose mother tongue = L1 and those students without a language section (SWALs) for whom the L1 is not their mother tongue.

- Alternative weighting models may be considered for the European Baccalaureate (Science) - see Recommendation 3.5.4.

3.5.6 Recommendation

We recommend a review of the weighting proposed in the European Baccalaureate Working Group:

- *The proposed reduction of weighting for the Preliminary Mark.*
- *The weighting of written examinations.*
- *The weighting of L1/L2 oral examinations relative to elective subjects.*
- *The combined weighting of L1 written and oral examinations.*

3.5.6.1 The recommendations of the European Baccalaureate Working Group are expressed in terms of making a change to the European Baccalaureate to "simplify organisation and to cut costs" but we believe that any changes should be seen primarily as ways to enhance the recognition

[31] Recommendations of DE parents of Brussels 111 re: proposals for the reform of the European Baccalaureate.

and value of the achievements of a student obtaining the European Baccalaureate. We believe, however, that our recommendations do contribute to organisational efficiency and cost-effectiveness.

3.6 Setting and production of assessment tasks in the European Baccalaureate

3.6.1 Overview

The **practicality** of an assessment is an important consideration. Assessments that are administratively and logistically straightforward are most likely to deliver on the goals of validity and reliability.

Two key elements of the operational context of an assessment concern the 'inputs' – the design and delivery of assessment tasks and the 'outputs' - the management and standardisation of the marking and reporting process. In this section we review the operational context relating to the production of assessment tasks, including both written papers and oral assessments. We understand that this is regarded as a particular focus of concern by those involved with the European Baccalaureate in terms both of current practicality and the ability of the system to scale up to meet expansion needs.

3.6.2 Written question papers

Our discussions with a wide range of stakeholders have given us a very helpful framework in understanding the production processes associated with the question papers for the European Baccalaureate. We understand it follows a cycle which begins in July/August of each year and continues until March the following year as follows:

Bac Unit sends out to European Schools a letter with subjects for which they have to prepare proposals: Bac Unit prepares and send out templates to be used (July-August).
Teachers prepare sample papers/questions. These are sent by the Directors to the Inspectors, the experts and Central Office. (August – mid-November).
Proposals, by subject, are put by schools on computer networks accessible by originating school and the Bac Unit (November)
Bac Unit moves each proposal onto the Learning Gateway by subject and provides access to Inspectors and Experts. (Access only to Inspectors and Experts for a particular subject) (November)
Definitive versions of the papers are prepared by the Inspectors and the experts. Translations of papers are prepared. At least one reserve paper is produced (December - February).
Papers are submitted to the Chairman of the European Baccalaureate for approval (March).
Final proof reading of papers under the responsibility of the Inspectors. Where necessary translations are completed (April).
Printing and reproduction and dispatch of the papers to the schools (May).
Written examinations are held and internal and external marking takes place (June).
Examinations are held (June - July).
Final grades are determined and deliberations take place (July).

3.6.2.1 Timeline of production

Some members of the evaluation team considered the process of designing the papers to be cumbersome and over-protracted. Those of the Cambridge team closely involved with the question paper setting process (where a question paper typically is constructed over an 18 month interval) did not necessarily share this view. However, the Evaluation team, as a whole, is agreed that further consideration of the question paper setting process is required.

The core issue is not the absolute number of stages or the timescale involved, but whether the system as a whole considers that the process of production is congruent with "fitness for purpose". The current processes, involving a number of participants and transactions, appear to be highly dependent on the positive, informal relationships which have built up between the different groups involved; the prospect of growth may have a negative impact on the procedure which the system as a whole is unable to control.

The establishment of the Bac Unit has done much to improve the tasks associated with setting assessments and improving the production values of question papers. It provides a co-ordination function and handles a large number of papers. In 2007-08 the Unit managed 153 different papers and 107 reserve papers, together with their associated mark schemes. The practice of working with teachers' proposals for question papers is a complicating factor as there are many more papers in the system prior to finalisation than is necessary. At an early stage there may be anything up to five times more question papers in circulation than the number needed. While many observers comment on improved production quality, nevertheless, the current administrative load, the volume of throughput at a busy time and the absence in part of syllabi aims, assessment objectives and specifications grids creates a vulnerability of the system which might become a concern given the further expansion of the European Baccalaureate. The observations of the Chairman of the 2004 Report continue to be to some extent true: "the uniformity of the format and of the level of the various papers examined is not the product of explicit and formal work involving common docimological planning on the Inspectors part but in fact results from the positive conversational climate which reigns among them. For that reason the tendency towards uniformity in the format and level of the written papers follows almost naturally from a common vision and participation in the didactic life of the schools, but is not the product of explicit definitions of protocols and of explicitly shared docimological criteria." [32]

3.6.2.2 Teachers' involvement in question paper production

A distinguishing feature of the European Baccalaureate approach to question paper production is the involvement of teachers as part of their teaching contract. This can be commended in so far as it helps to ensure that teaching, learning and assessment are integrated in the curriculum process; it affords teachers a unique opportunity to exercise the full range of their professional skills and can form an important dimension of their own professional learning.

There are, however, potential risks in teachers' involvement in producing the final assessment tasks and, unless appropriate checks are put in place, the standards and quality of the examination can be compromised and public confidence in the system subsequently eroded.

[32] Report of the Chairman of the 2004 European Baccalaureate Examining Board Ref.: 2004–D–3110–en-1.

This becomes particularly important when European Baccalaureate expansion is contemplated.

The size of entry for some of the subjects available in almost all schools can mitigate issues of credibility related to the involvement of teachers. However, small entry subjects available only in some schools cannot take advantage of this factor. That it is a current problem is demonstrated by Article 6.3.3:[33] "should an examination be organised in only one school, it is particularly important for the Committee of Experts to find ways of guaranteeing the credibility of the examination. When possible, even if a subject is taught in only one school, at least two schools should submit proposals."

It is difficult to see how this is designed to work in a situation where only one school has first hand experience of teaching the subject.

Other assessment-related concerns which stem from teachers' involvement in the development of question papers include:

a. Incomplete sampling of syllabus

While Inspectors were robust in their assertion that teachers' proposals for question papers were not accepted without modification it appeared to be the case that there was considerable subject-to-subject variation. The L2 English Examiners who were involved in setting question papers considered that their draft papers were not subject to significant change by Inspectors. In another context, the pool of papers in a subject of small take-up, for example Polish, inevitably constrains the degree of choice available.

In determining the final version of a question paper it can be difficult to deviate from the original samples when devising final versions of the paper. The comment on Economics in the 2003 Report of the Examiners is a telling one: "It was disappointing that there were no questions requiring calculations but there were none suitable submitted by the schools this year". It is understandable that teachers' views of areas of the syllabus that their students might find difficult might produce a tendency towards topic or skill avoidance, which needs to be strongly countered in the production process.

The requirement to sample the syllabus as a whole involves not just syllabus content. It is also necessary to do this across a number of dimensions: to ensure that a pre-defined taxonomy of **assessment objectives** is achieved (weighting between knowledge, skills, application, analysis and evaluation) and that the question paper complies with a pre-ordained **assessment structure.**

3.6.2.2a Recommendation

If teachers are to remain involved in the question paper setting process, a wide range of materials should be provided to them and sufficient time should be given to enable discussion of student performance in previous sessions, the overall European Baccalaureate and subject goals and assessment objectives, and the Chief Examiners' analysis of the session.

[33] Arrangements for Implementing the Regulations for the European Baccalaureate 2007-08.

b. Predictability of assessment tasks

The engagement of teachers in setting assessments can lead to a concern that teachers (and students) are able to predict with reasonable certainty what the examination will contain, and that teachers might share this information across the system. We received no evidence to support this contention. Indeed the phenomenon of 'leakage' is more likely to happen in cases where teachers and schools feel that they are being held accountable for the performance of their pupils in public examinations rather than in the 'educational family' established by the European Schools.

Nevertheless, the perception of an 'open' system can lead to a lack of confidence and again this is a factor that will require careful consideration in the context of expansion.

Teachers, by virtue of their involvement at the first stage in the design of the papers, can exert significant influence over an unchanging format in the nature of the examination from year to year. Predictability in format could, itself, be compounded by the fact that teachers, unused to the setting process, rely too heavily on previous question paper and assessment models for guidance. The perception of predictability in question papers may in turn impact on the overall credibility of the system.

One further impact of 'predictability' lies in the avoidance of unfamiliar contexts in question papers. Experienced question-paper setters become very used to developing an unfamiliar context or application to test the extent to which a concept has been mastered. Our analysis of European Baccalaureate question papers against national systems found too little evidence of this phenomenon. There is a tendency for teachers, with a strong understanding of students' learning styles to introduce 'difficulty' through the assessment of 'hard' concepts rather than 'easy' concepts applied in challenging, new contexts.

This phenomenon applies even for Advanced Mathematics where the 2007 Report to the Board of Governors observes: "The questions are purely technical questions; only an application of rules without any mathematical thinking or reasoning".[34]

A further dimension to the issue of predictability is a matter of concern, not because of the actions of individuals in the process but because of the operation of the system as a whole. There are many aspects of a question paper which should avoid 'surprise' to the test-taker. Question formats, styles and sequencing should fall within a recognisable pattern to avoid undue stress on the test-taker.

Question-spotting, however, leads to a situation in which the overall curriculum is narrowed to what is likely to be on the question paper producing a negative washback on the curriculum as a whole. Our interviews with Examiners,[35] supported by our comparability studies[36] indicated to us that external question papers for the European Baccalaureate could be regarded as more predictable than those for other examination systems. The steer towards the assessment of the year 7 curriculum often accentuated by the structural requirements of particular syllabuses, can make this difficult to avoid.

---

[34] Report of the Chairman of the 2007 European Baccalaureate Examining Board 2007-D-2010-en-1.

[35] Interviews conducted in June 2008 during the Residential marking exercise in Brussels as part of the External Evaluation of the European Baccalaureate.

[36] See Chapter 1 of this report.

Whereas, Article 4.1 European Baccalaureate [37] is no doubt designed to ensure that "new" content, i.e. that not previously assessed internally, is assessed in the final examination, it can have the result of narrowing the curriculum unduly. For example, the subdivisions necessary to cover distinct areas of a subject, such as Inorganic and Organic Chemistry, were felt by some examiners to produce a paper that had greater predictability than in other examination systems.

The assessment at the end of year 7 is the conclusion of a two-year programme. The requirement that the tasks should be steered toward the year 7 curriculum might be interpreted by teachers in setting question papers too strongly. We recommend that the steer given in Article 4.1 should be one of encouraging teachers to set questions for students which bring together different elements of a course in a way which provides the opportunity to demonstrate mastery of learning.

3.6.2.2b Recommendation

The sign off process by the Chairman of the Examination Board should involve a requirement that a full check has been conducted against previous years' question papers to avoid a tendency towards the predictability of question paper formats.

Consideration might be given to the use of a training workshop on item-writing for teachers to ensure that items within a question paper provide appropriate challenge.

We recommend that the advice to steer the final assessment to year 7 topics be clarified in a way which makes clear that the question paper should sample across the whole programme of study. We also recommend that question styles encourage the synthesis of topics taken from across the two year programme of study.

If teachers are to continue to be involved in the question paper setting process we recommend that when a syllabus is revised, or a new one introduced, specimen papers are produced, perhaps outsourcing the work to provide assessment exemplars for teaching and learning purposes in advance of the first set of question papers which teachers would be asked to draft.

c. Impact ('washback') on curriculum goals

The design of a question paper has an inevitable influence on the curriculum as a whole as teachers prepare their students for examination. The involvement of teachers in the question paper setting process can have an impact on the achievement of curriculum goals, if the question paper does not fully reflect curriculum aims and objectives. It can also lead to a situation in which teachers become resistant to syllabus revision. In fact the current mechanism of question paper setting makes no provision for significant review in syllabus design. Established international practice, for example, adopted by the International Baccalaureate, by University of Cambridge International Examinations and by examination boards around the world accompanies each syllabus revision with specimen papers and mark schemes to familiarise teachers with new or changed requirements. The issue arises – if

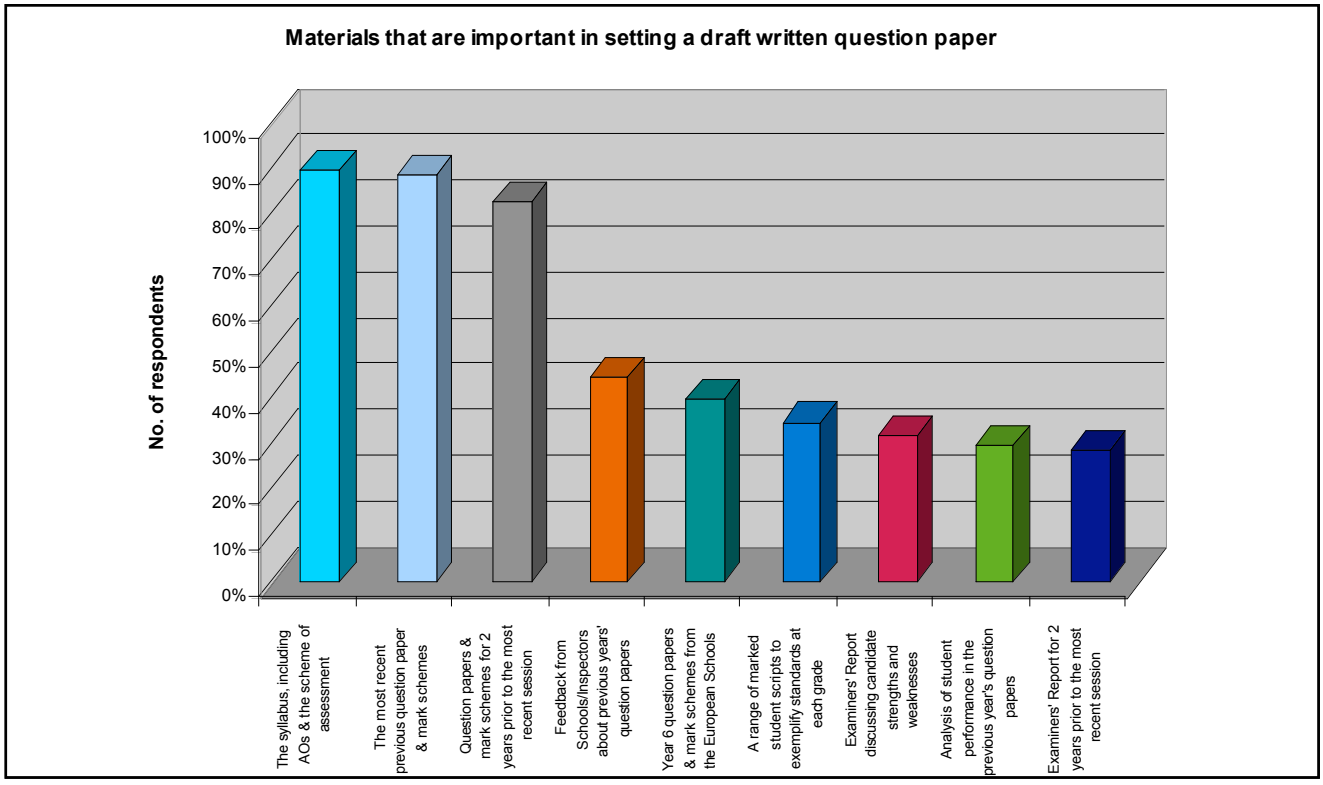
[37] Arrangements for Implementing the Regulations for the European Baccalaureate Applicable for the 2008-09 School year Ref.:2008-D-58-EN, Article 4.1.

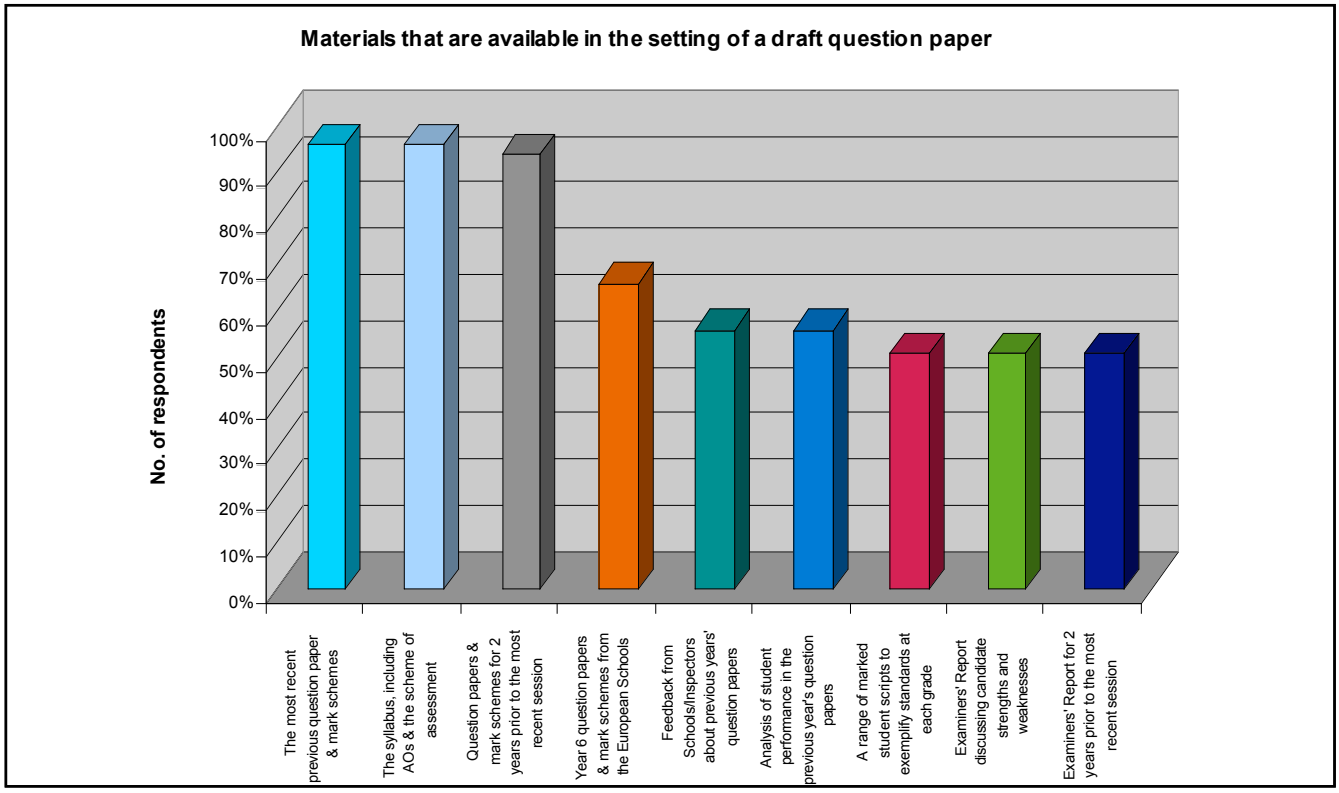
such practice were to be adopted by the European Schools where would the responsibility for question paper production reside?

3.6.2.3 The Setting Process

During the setting process a wide range of reference and support materials need to be available to setters. In reviewing teachers' involvement in the question paper setting process we carried out a brief survey in late October 2008. 215 teachers responded. Again, high levels of experience were noted. 80% of respondents had worked in European Schools for more than two years and the overwhelming majority had previously taught in their home country for more than ten years (57.2%). 86% were on secondment contracts and 14% had been recruited as a result of direct application to a European School. Teachers were asked about the materials they felt were important for them to have available in their task of question paper setting.

The syllabus (including assessment objectives and the scheme of assessment) was regarded as the most important reference document (95%) together with the previous year's question paper (95%) and those for the two years previous to that (94%). Encouragingly, these materials were overwhelmingly available; a little less encouraging is the significant drop in availability or importance to teachers of a range of other materials which might typically be regarded as crucial to assessment best practice. Around 50% of teachers did not consider it particularly important to refer to documents other than previous question papers and the syllabus. Nor were such materials available to them.



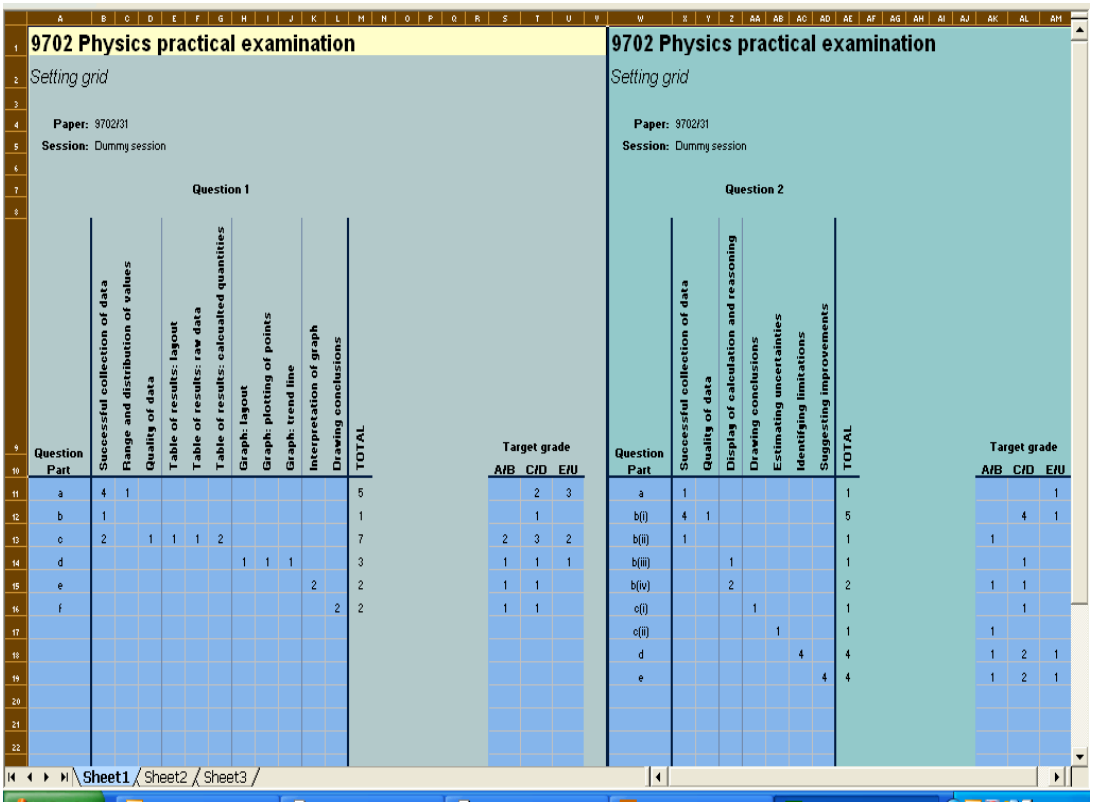


3.6.2.3 Recommendation

Those responsible for setting papers must increase their use of a wide range of research materials to reflect best practice in their construction of question papers.

3.6.3 (a) The use of a specification grid

To ensure that a question paper is fit for purpose in terms of its assessment aims and objectives it is considered good practice to work within a specification grid. The diagram below provides one such example of a specification grid in which a question's link with defined assessment objectives and assessment content is established and in which the author is required to identify the applicability of the question for all ability ranges within scope.



9702 Physics practical examination

Setting grid

Paper: 9702/31

Session: Dummy session

Question 1

Question Part	Successful collection of data	Range and distribution of values	Quality of data	Table of results: layout	Table of results: raw data	Table of results: calculated quantities	Graph: layout	Graph: plotting of points	Graph: trend line	Interpretation of graph	Drawing conclusions	TOTAL	Target grade A/B	Target grade C/D	Target grade E/U
a	4	1										5		2	3
b	1											1		1	
c	2			1	1	1	2					7	2	3	2
d								1	1	1		3	1	1	1
e											2	2	1	1	
f												2	2	1	1	

9702 Physics practical examination

Setting grid

Paper: 9702/31

Session: Dummy session

Question 2

Question Part	Successful collection of data	Quality of data	Display of calculation and reasoning	Drawing conclusions	Estimating uncertainties	Identifying limitations	Suggesting improvements	TOTAL	Target grade A/B	Target grade C/D	Target grade E/U
a	1							1			1
b(i)	4	1						5		4	1
b(ii)	1							1	1		
b(iii)			1					1		1	
b(iv)			2					2	1	1	
c(i)				1				1		1	
c(ii)					1			1	1		
d						4		4	1	2	1
e							4	4	1	2	1

Sheet1 / Sheet2 / Sheet3

3.6.3 a. Recommendation

The European Schools should consider the use of such specification grids in setting assessment tasks that representatively sample syllabus content and assessment objectives and that comply with the assessment structure.

(b) Training to support the setting process

Devising examination questions and papers of high quality requires special skills and expertise as well as competence in the subject concerned. It would be unwise to assume that teachers automatically acquire this expertise; they need regular updating in the skills required. It is not clear that they receive appropriate training in the development of examination techniques or are given precise guidelines on the preparation of questions.

3.6.3 b. Recommendation

We recommend that occasional training be provided for teachers specifically designed to develop the skills of item writing and question paper construction. Their involvement in internal assessment makes this necessary whatever their role in setting questions for European Baccalaureate examinations.

3.6.4 Setting question papers at the level of items or complete question papers

There are two schools of thought about the way in which question papers are constructed. One takes the view that a question paper can be constructed from item-level up (i.e. individual questions) working against a specification grid to ensure appropriate sampling of the syllabus

content. The second approach emphasises that a question paper should be constructed as a whole – one setter taking responsibility for the internal 'fit' of the assessment produced. Teachers for the European Baccalaureate are asked to supply papers as a whole, adopting the latter approach. However, the encouragement is given to subject experts to de-construct papers to ensure that a composite 'best' paper is produced, i.e. an amalgam is produced from sections of each teacher's work. The overall approach, therefore, offers a 'mixed economy' in which the benefit of neither method is achieved.

We consider, in the context of the overall importance given to the role of teachers as assessors in European Schools that teachers should continue to be involved in setting, but that the emphasis moves to one of drafting sections of a paper rather than the whole paper and that a minimum of three variants of each section are commissioned. This gives the final setter sufficient material from which to construct an appropriate assessment.

3.6.4 Recommendation

We recommend that teachers continue to be involved in drafting assessments for the final examinations of the European. However, we propose that their role moves to one in which individual teachers are commissioned to produce defined sections of an assessment which sample a syllabus, assessment objectives and scheme of assessment in a manner identified by a specification grid.

3.6.5 Involvement of external experts in setting question papers

The process by which initial submissions from teachers are compiled into a question paper is currently complex and likely to be more so with a further expansion of the European Baccalaureate and the need to introduce more questions. For this reason we also agree with the recommendations of the Working Group that external experts should be contracted to have a role in the question paper setting process.

3.6.5 Recommendation

We agree with the recommendations of the Working Group that external experts should be contracted to have a role in the question paper setting process. We consider that the specification of their competency should also require, in addition to the features identified by the working group, experience of question paper setting.

3.6.6 The Production Process

Well over 100 papers are processed in a short timescale by the Inspectors and the Bac Unit. In addition to comments that we have made in the preceding sections we add some observations on practicalities of the process.

3.6.6.1 Use of a single written examination paper

Typically, a single written paper for each subject is used. The 'one chance' nature of this can place a candidate under considerable pressure. In order to reduce the impact of any temporary factors on a candidate's performance it is often thought desirable to structure an examination into more than one component. In addition, the fact that there is only one

examination paper per subject means that risks attendant on transporting scripts become greater, which becomes a relevant issue if the postal transporting of scripts is considered.

3.6.6.1 Recommendation

It may be appropriate to review the practice of providing one written examination in a subject. Two examinations of shorter length deliver both assessment and administrative advantages.

3.6.6.2 Question paper error

The process of expert panel review of question papers is a rigorous one but some errors can occur. For example, there is a danger of some topics not being on the syllabus – delineations of subject domains show variation across different national systems: the definition of a Biology syllabus in Italy would contain items more usually covered in a Physical Geography or Geology paper in the UK and this phenomenon can lead to items being included within an examination paper that fall outside the defined curriculum. Such differences are difficult for any one individual to detect reliably. Thus, although there is a process of final sign-off by the Chairman of the European Baccalaureate Examination Board it would be unreasonable to assume that such errors can be detected at this stage. We refer once again to our proposal that the use of a specification grid is introduced into the setting of question papers. The quality assurance role of the Chairman would then be one of checking that a specification grid, together with the mark scheme, had been correctly completed.

This would ensure that a reconciliation of items (individual questions) and syllabus has been carried out.

3.6.6.2 Recommendation

The process of the March sign-off of question papers by the Chairman should be accompanied by a review of evidence of a compliance check against syllabus documentation for each subject.

3.6.7 Print production

The final production of the papers is organised by the Bac Unit. Interviewees acknowledged that considerable improvements have been achieved in the production of papers in recent years. This is most noticeable in the precision and clarity of the graphics and the use of colour such as in maps for Geography. Teachers' involvement in question paper proposals included the submission of 'texts, figures and diagrams' for use in the examination. This raises the question of obtaining copyright for usage. We are not clear at what point copyright permissions are acquired. In the case of diagrams, it is sometimes useful to refer to databases of scientific and mathematical illustrations. A copyright-free CD-Rom of scientific graphics is available from Cambridge Assessment and may save some production time for teachers.

3.6.7 Recommendation

We recommend review of the processes for obtaining copyright for texts and illustrations used within question papers. It may be of interest to review the applicability of materials contained on the CD-ROM of scientific graphics available from Cambridge Assessment.

3.6.8 Currently, the Bac Unit outsources some printing. Secure printers offer a highly specified service and consideration might be given to a more extended use of secure print services. We return to these issues in Chapter 5 of this report.

3.6.8 Recommendation

Consideration might be given to the extended use of secure printer services.

3.7 Current processes of marking and review

3.7.1 Overview

Taking the European Baccalaureate examination as a whole the key issue concerns marking reliability in a way that includes both internal and external assessments. The European Baccalaureate operates with a high percentage of internal assessment. The quality of teaching and learning points to the benefits of this approach and is observed annually by external reviewers.

In the context of marking quality the scope of interest – from the first mark given by a teacher to a student in Year 6 - includes all marking. The system needs to be able to provide reassurance that each instance of continuous assessment is marked to an agreed standard between different teachers in the same school and in different schools, across language sections and across different subjects, in order to ensure the equity of the system. In a national system there are inevitably teacher-to-teacher variations but standardisation exercises can address this tendency at the beginning of a marking cycle; a shared national consensus minimises variability and a moderation process can be established to bring everyone into line.

A Swedish subject teacher will share many implicit assumptions about standards with other Swedish teachers of that subject which derive from their common educational heritage and professional training. Across the different systems of European education however, such implicit assumptions do not hold true.

In systems with a large weighing of internal assessment it is important that teachers have a strong and stable idea of the standard to be applied and that this is reliably applied across all institutions. This requires robust strategies across national systems in the European Schools' system.

3.7.2 Teachers' marking

In the case of the European Schools, operating in a trans-national context, there is a difficulty which follows from a recruitment pattern which brings together those teachers seconded from a national system and those who are locally recruited. On the evidence of survey

responses[38] a significant proportion of staff in European Schools are recruited locally and contracted on a part time basis. Directors report variously that such teachers make up 25% to 50% of the full-time equivalent teaching establishment. The average of all responses in this respect was 31.8%. Many of these are involved in teaching in years 6 and 7. Such teachers may be excellent in terms of their professional expertise but nevertheless create vulnerability in terms of their connectivity with the European Schools' system.

To add to this, the turnover of teachers runs at an average level of 10-25%. These two factors taken together indicate the difficulty which Directors face in ensuring a coherent European Baccalaureate standard in years 6 and 7.

The subjects that such locally recruited teachers provide are spread across the curriculum. One school responded that 'we try to avoid giving lessons for the locally recruited teachers in the 6th and 7th year of secondary' but in a number of other cases teachers with a wide range of standards are being recruited into the system on an ongoing basis.

The process of ensuring that all teachers share a common view of subject standards needs to be rigorously established to safeguard levels of internal assessment.[39]

In respect of the strategies which might be employed in schools to ensure that teachers share a common view of standards, Directors' responses to the survey echoed those of the teachers that there were low levels of either generic induction programmes for teachers (75% of Directors' responded that such programmes rarely or never took place) or generic professional development (which 37.5% of Directors considered rarely or never took place).

There is a need therefore to embark on establishing processes to bring standards into line within an overall European Baccalaureate standard. The development of a strong and consistent European Baccalaureate standard for both internal and external assessments is a key opportunity and essential precursor of any further expansion of the system.

3.7.2 Recommendation

Arrangements should be put in place to ensure that markers are standardised before they begin marking.

3.7.3 We also agree with the recommendation of the European Baccalaureate Working Group – "Given that there does not appear to be agreed criteria for the internal components, how can it be established that similar standards are being applied across the different schools? It would appear reasonable that the internal components should meet the same standards and stand up to the same scrutiny as the external components. As a first step in this process, the distribution of internal marks within and across schools could be analysed and also correlated with those of the written and the components."[40]

The introduction in the 2008 Report of an analysis of marks obtained in written, oral and preliminary components of the BAC begins this process. The Report on the Joint Inspector to the European School at Karlsruhe is also useful in this respect. It contains an analysis of Preliminary marks against students' subsequent written examination marks:

[38] Survey of Directors of European Schools, conducted July 2008.

[39] See Section 3.3.2.

[40] Report of the Chairman of the Baccalaureate 2003 Ref.:2003-D-3610-en-2.

"…in the Baccalaureate in 2006 the students achieved an average of 7.0 in the written examinations in history, while the average B mark in S4-S7 was 7.9; for Geography the average in the 2006 Baccalaureate written examinations was 7.3 while the average B mark in S4-S7 was 7.7. So, ultimately, the results in the Baccalaureate were somewhat lower than the teachers expected".[41] It is important that teachers and Directors are provided with the opportunity to discuss the observations of the Joint Inspector visits.

A Report on the European Baccalaureate in 2007 notes that in general the scores given by the teachers are too high regarding the quality of work and suggests that formal marking conferences are needed to review the discrepancy in assessment. We concur with this proposal.

3.7.3 Recommendation

A formal analysis of the discrepancy between teachers' and external examiners' marking should be carried out. Formal marking conferences should review discrepancy of this nature.

3.7.4 Double marking

The structure of the European Baccalaureate might be described as one of internal assessment which is then validated by an external assessment exercise. However, the involvement of teachers in the external assessment complicates such description.

An established part of the European Baccalaureate is that teachers complete the first marking of scripts immediately after the examination. The involvement of teachers at this stage requires further consideration. In terms of a formal model it is not clear what purpose this marking provides. It may variously be defined as:

- A first marking of equal value to the second marking conducted in the external marking exercise
- A pre-emptive marking to flag up any evidence of mis-marking at the second stage
- A further internal assessment.

This last definition would appear to be the one in common usage providing, for the same examination, both an internal and an external assessment score.

The marking conducted by teachers immediately after the examination falls short of established assessment practice on a number of counts:

i) teachers have not been 'standardised' at this stage, i.e. they have had no training in the standards to be expected on the examination in this session.

ii) the speed at which teachers are required to mark written papers can compromise thinking time and internalisation of the mark scheme. Their practice, as it were, is on live scripts.

[41] Final Report of a Joint Inspection/ES, Karlsruhe, Ref 2007-D0137-en-3, p10.

iii) access to the mark-scheme accompanying the question paper prior to the examination is an unusual procedure and one which might need to be reviewed with the further expansion of the European Baccalaureate.[42]

iv) teachers' views of 'good' candidates and the established ranking between students may be expected to interfere with objective marking of the evidence before them in student scripts.

In the case of the European Baccalaureate, the implicit assumption is that the teacher is as effective a marker as the external marker. We did not notice any mechanism to challenge this assumption, nor any to demonstrate its accuracy. The external examiners referred to the fact that "teachers give higher marks" but as no marks are recorded on the papers it is difficult to see what script characteristics are being marked differently.

The practice of allowing a 20% tolerance in marking between the first and second correction might be considered a much greater differential than would be considered appropriate in other high-stakes examinations. Given that the marking predominantly uses a range of 60%-100%, averaging over a 20% difference is extremely significant. For those universities who have stipulated admission criteria in terms of "an average of 7.5 and an 8 in a particular subject" the fact that this has been achieved by averaging a higher internally-assessed mark with the external assessment is of material interest.

3.7.4 Recommendation

The process which allows a 20% tolerance between two marks should be reviewed.

3.7.5 Marking standardisation models

In criterion-based assessments that call for the exercise of professional judgement there is a need for the marker to be 'standardised' – i.e. for a test exercise to demonstrate that the standard has been internalised in marking. All markers are required to take part in an exercise of this nature with the exception of the most expert marker (the most senior in a hierarchy of markers).

In the Irish, UK, International and French practices most familiar to the Evaluation Team, the practice of ensuring that each marking session begins with a training exercise is common.

A test exercise is provided in most systems. Script sampling and discussion of the application of the mark scheme is common. Examiner trial marking provides assurance that the correct standard is being applied.

The type of examiners selected to take part in the residential marking exercise for European Baccalaureate scripts are those with considerable marking experience. Nevertheless, the fact that they are most typically recruited by Inspectors means that much of this experience will have been obtained in a national system, and national referent points will be strongly

[42] We note in the 2008 Report of the Chairman to the Board of Governors that the practice of sending mark schemes and sample answers at the same time as the question papers will be revised. From 2009 the intention is to send such materials electronically and only on the day of examination. This is a sensible revision to existing practice.

developed. For this reason, we regard it as essential that a standardisation process is developed.

This Report of 2005 refers to an "interesting experiment in harmonisation" and in so doing points to the important assessment principle of standardisation. The Chairman observed the practice by which the external examiners of L2 and L3 first consulted one another on the award of detailed marks for different parts of question responses. Then they corrected two photocopied scripts separately. Finally comparison of the results enabled one or other of the examiners to realise where their marking was too strict or lenient.[43]

The Chairman goes on to say that whilst such experiments should be encouraged it should not become 'general practice' because subjects such as Science operate with a mark scheme which is more prescriptive. However, we would comment that at the European Baccalaureate level even in Science a mark-scheme requires interpretation and application and we would advise that the approach does become general practice. The fact that a nationally-skewed profile of results can be observed for the Sciences as much as for languages indicates the extent to which this would optimise practice.

3.7.5 Recommendation

We recommend the introduction of a standardisation process to ensure that external examiners across different subjects mark at the appropriate level before they begin their marking in each session.

3.7.6 A proposal for change: a teacher's role in rank ordering students

One option which might be considered is that the role of the first corrector be re-defined to one of producing a rank order of students. Current assessment research indicates that teachers are very well equipped to rank their students against established criteria. On a smaller scale and in a different context it seems to describe examiners and teachers' practice in respect of oral assessments where for any three students a ranking is produced and then extended to encompass the performance of the next three students and the next three and so on. This is a very effective use of human judgement. The rank ordering of student scripts and performances avoids the allocation of marks and this may be considered a positive development because:

- It recognises that teachers understand very well the relative ability of their students – that X is better than Y but not as strong as Z.
- It avoids the despatch of a mark scheme to a school at the same time as question paper. Current practice might be regarded as a threat to the security of the administration.[44]
- It removes the inequalities created by one corrector (in Brussels) having the opportunity either informally to talk to other examiners or formally to be standardised in the application of the mark scheme and the other corrector not having the equivalent opportunity.

[43] Report of the Chairman of the 2005 Baccalaureate Examining Board Ref.:2005-D-3010-en-2.

[44] We understand from the 2008 Annual Report on the European Baccalaureate that the intention is to discontinue this practice with effect from 2009.

- It avoids the position articulated by one of the Evaluation Team in which "the truth is never the average of two errors".

Such a proposal also recognises the concern of a number of stakeholders that a reduction in the volume of internal assessment would materially affect the validity of the qualification. "The leaving certificate of a secondary school should not be restricted to reporting scores obtained during the final examinations (which give an indication of the pupil's knowledge at that given moment in time) but should also reflect a pupil's global competences and performances demonstrated during the school year through balanced assessment. Furthermore, a final mark based exclusively on centralised examinations does not reflect competences which students need to acquire before pursuing university studies."[45]

Using a teachers' rank order of a group of students can take account of class participation in the second semester of year 7 as well as the script outcomes.

The Senior Examiner would be able to review external examination marks which did not follow the rank order of subjects, thus continuing to provide the quality assurance of marking.

3.7.6 Recommendation

Consideration might be given to a variation of the current 'double marking' system. Our proposal is that teachers would be asked to draw up a student rank order before sending scripts to Brussels. The rank order would be sent to Brussels sealed. A double external marking would be carried out for any students for whom the rank ordering was interrupted.

We recommend that all markers for internal and external assessments of the European Baccalaureate undergo a process of 'standardisation'.

3.7.7 Residential marking exercise

We are aware of many current concerns about the practice of residential marking in Brussels. The focus of such concern is that the practice is very expensive and applies equally to those subjects where many examiners need to meet together to discuss a common application of the standard and those where the subject has so few entries that only one examiner will be appointed.

During the two days which the Evaluation Team spent in Brussels we noted that the residential marking exercise brought together not only the examiners involved in marking but also the Inspectors with lead subject responsibility, the School Directors and the staff of the Bac Unit. From the assessment perspective the residential marking in Brussels fulfils a number of functions. Inspectors and experts from different subject areas come together at an important time of the year to review standards and outcomes, and examiners have the opportunity to meet each other and synthesise national approaches. Administrative matters are dealt with. These are all valuable activities. However, the multi-functional nature of the marking exercise can mean that some potential benefits of an expensive residential marking process are not necessarily realised.

One of the values of residential marking of the type that was conducted in Brussels is that there can be exhaustive discussions between examiners on the features of a common

[45] Interparents Memorandum, 2008, p4.

photocopied script to ensure that a common application of the mark schemes is being conducted. This did not routinely happen. Examiners with very small marking loads had completed their tasks quickly, and would in all likelihood be the single examiner for a subject. We did not observe consistent practices of trial marking and marking evaluation by team leaders, although informal relationships between examiners led to discussions about different national approaches in a subject.

Processes of cross-moderation by which outcomes across all subjects, across schools and within clusters of cognate subjects (all Sciences, all Languages) etc were not formally scheduled. We return to the subject of residential marking in Section 5.10.9. We consider it would be possible, with no loss of quality, to review this practice.

Although much work has been done in the development of marking schemes in European Baccalaureate assessment there remain other areas of concern in comparison with national practice:

- Although a marking scheme is provided there are few quality checks in place to ensure its use and effective interpretation.
- The conditions of marking – time and environmental issues – can mean that marking is conducted too quickly.
- There is no concept of a 'jury' (France) or Grade Review (UK).

In France, there is the potential for escalation to a 'jury'; in Cambridge the practice of reviewing candidates who meet certain criteria and who might be 'at risk' of an incorrect mark leads to a process of further professional judgement. The trigger for escalation in the residential marking exercise is plus 20% disagreement between two marks. We did not observe any other sampling of either teacher or examiner marking.

An external jury has the responsibility of checking the quality of each examiner's marking. In the case of France the trigger for this is when an examiner's results are too different from others (average and distribution). The jury has the final authority over the mark to be issued. In England and Wales candidates are re-marked when their scripts are either at a threshold or too different in outcome to others for that candidate. It would be possible for procedures in the European Baccalaureate to double mark only those scripts where the final outcome is significantly different to that achieved in the Preliminary Marks.

In cases where an examiner's marking load justifies the use of statistical analysis, each examiner's distribution profile of results should be reviewed by the lead examiner.

To summarise our views on a topic which has frequently been the focus of attention in the Chairman's Reports, in Reform Proposals and in the Annual Report - the residential nature of the marking:

- There is no particular need to co-locate all examiners.
- Standards discussions most frequently take place on a face-to-face basis and due consideration must be given to the way in which this would be done without such a meeting. An intranet and telephone link between the inspector and examiners should be established.

- Many examination boards are moving towards 'virtual' standardisation approaches, particularly using digitised scripts and on-line marking. The move away from a residential marking exercise requires detailed consideration and it might be sensible to make the move in a number of stages.

Stage i: Subjects currently marked by a single examiner

Stage ii: Subjects with a team of experienced examiners, preferably for those examinations for which there is a tightly specified mark scheme.

Stage iii: Panels of experienced examiners with less prescriptive mark schemes.

Stage iv: Whether or not the process is rolled out fully to panels of 'new' examiners in subjects with less prescriptive mark schemes will be a matter for further consideration once earlier stages have been evaluated.

3.7.7 Recommendation

We recommend a phased move away from residential marking. From the assessment perspective, there is not strong evidence that the process as currently constituted yields benefits that could not be otherwise replicated.

3.8 Oral assessment

3.8.1 Overview

Oral assessment forms a very important part of a student's programme on the European Baccalaureate. In this section we:

- make some observations about current practice
- review the proposals for reform put forward by the European Baccalaureate Working Group
- make some recommendations for ongoing practice which might be appropriate in the context of the expansion of the European Baccalaureate.

3.8.2 Observations on current practice

The management of oral examinations can be administratively time-consuming and complex. The logistical focus entailed in their management can overshadow issues relating to the assessment objectives of the exercise, i.e., in what ways the assessment of the candidate enhances the validity of the examination as a whole. As with written examinations it is important that an assessment is "fit for purpose".

The 2008 Report of the Survey of Schools relating to the conduct of oral assessment provides a comprehensive commentary on the current situation. 314 questionnaires were completed covering L1, L2 and L3 orals, Biology and the elective subject of choice. Good results were observed across a range of administrative areas – the examination was completed within appropriate time allocations; printed stimulus materials were almost entirely accurate and the different components of the assessment were appropriately structured. Less high scores are

reported for “candidates are given sufficient time and opportunity to express themselves” (75%) and for the use of written assessment criteria by both the internal and external examiners (60% yes; 40% no).

This represents an encouraging increase on the 19% using written assessment criteria in 2007 but it is lower than desirable. The recording of greater participation by the internal examiner (at 17% against 12% for external examiner) throws some doubt on who is seen to ‘own’ this examination. Is this an internal assessment moderated by an external examiner or an assessment with double marking of equal value?

We commend the target that in 2009 written criteria will be available for the assessment of the pupil’s performance in all subjects. We also commend the practice of gaining feedback on the conduct of assessments each year.

3.8.3 Opportunities for improvement

3.8.3.1 Timetabling of oral assessments

While the evaluator was impressed by the quality of the interaction and professionalism witnessed, the short timescale between the oral assessment and the publication of marks caused concern. The Evaluation Team were not able to attend many oral assessments but were present for one day’s assessment in Culham that coincided with the end of the assessment period.

Following deliberation by the school, students were being contacted that evening with their European Baccalaureate result. There was no opportunity for any review of marking across the system and across the outcomes of previous sessions. To make this point is not to assume that norm-referenced outcomes are required but rather to recognise that in the case of disputed judgement between an examiner and a teacher, there should be time for escalation and further consideration.

There should also be the opportunity for a review across the outcomes for all schools to enable consideration of the year-on-year trends and to ensure that L1/L2 written performance and Preliminary Mark performance was in line with putative oral outcomes.

3.8.3.1 Recommendation

We recommend a review of oral assessment timetabling so that there is the opportunity for a cross-moderation of standards before the issue of results. Our proposals in Section 5.10.5 discuss the way in which this can be done.

3.8.3.2 Design of oral assessment tasks

There are still some concerns about practical administration. The first of these relates to the design of assessment tasks for orals. The responsibility for the writing of oral questions falls to a subject teacher. There is no process of quality approval or moderation of the task. Re Article 6.4.3.2: “The examination questions will be proposed by the subject teacher”. This is a risk to both the validity and reliability of the system. The chairman of the 2005 Report observes: “I would assume that all the questions set by the class teacher are of an equal standard and of an equal degree of difficulty”. Apart from the fact this remains an

assumption, from the point of view of an assessment system, the key question goes further: is this the same standard as for other teachers in other schools, and for the previous session?

The question is not simply one of standard but of appropriate sampling and avoidance of predictability. Are the teachers' topics/questions appropriate and 'novel' enough to avoid becoming familiar to the test taker?

3.8.3.2 Recommendation

We recommend a review of the practice by which assessment tasks for oral assessment are designed by the student's teacher. Efficiencies and rigour might be increased by a move toward a centralised design of oral assessments which are provided for teachers in the same way as written examinations.

3.8.3.3 Number and nature of oral examinations

Currently candidates take four oral examinations: Language 1 (or Advanced Language 1), Language 2 (or Advanced Language 2), History, Geography or Philosophy (whether two or four-period options in each) and one other.

The European Baccalaureate Working Group has reviewed the range of subjects for which oral assessment is available. We agree with proposals for reform in this respect. The volumes of candidates taking assessment in subjects such as Economics (8), Physics (8), Chemistry (10) (Report on the European Baccalaureate 2004, p11) are too small to be able to make sound judgements of standard. The removal of the Advanced Mathematics oral also seems sound.

However, the aims and objectives of the syllabuses for History and Geography are so inextricably linked with a candidate's ability to communicate in their second language that we consider it important to continue the system of oral assessment for these subjects.

3.8.3.3 Recommendation

We agree with the reduction in subjects for which there are oral assessments. The removal of the oral in Advanced Mathematics seems sound. However, we consider that the aims and objectives of the syllabuses for History and Geography are so inextricably linked to the ability to communicate that we would recommend retention of oral assessment in these subjects.

3.8.3.4 Duration

An oral assessment is specified as being 20 minutes long. In the case of L1 and L2 examinations, it adds to the profile of a student's performance in Speaking/Writing. When an oral is taken in other subjects it forms the only assessment of a students' ability and achievements at that stage and concern must be raised about the ability of a 20 minute oral to sample effectively across the syllabus when the dominant mode of the written examinations is a three/four hour examination "For the four period Geography it would be advisable to extend the length of the oral examination to at least 30 minutes. It was very difficult to do justice to the greater depth and breadth of the syllabus in the 20 minutes allotted".

3.8.3.4 Recommendation

We recommend review of the duration of oral assessments. Whereas 20 minutes would seem to be appropriate for L1 and L2 assessments for which other performance exists from

written papers it would seem an inadequate duration for the assessment of History or Geography.

3.8.3.5 Oral assessment of extended project work

To refer back to our proposal in Section 2.5.2 that consideration be given to the introduction of an extended piece of individual research or project work, the assessment of this might take the form of the work itself and an oral assessment in connection with it.

3.8.3.5 Recommendation

Consideration may be given to the introduction of an oral assessment linked to a presentation which might have been completed as part of a cross-curriculum piece of work (see Section 2.5).

3.9 Aural assessment

It is accepted good practice that language assessments should sample across the four skills: reading, writing, listening and speaking. Although a candidate's listening skills are clearly sampled in the course of an oral assessment designed to assess both speaking and listening, there is scope for the development of an aural test standardised across all schools.

A listening comprehension test has the potential to add to the range of evidence available on candidate performance, may be externally marked and avoids heavy cost. Aural tests can be digitally recorded and despatched on CD-ROM to schools. Alternatively, although with less reliability, teachers can be asked to deliver the test in the examination room. While cost considerations may initially seem to be an issue for any other than the three working languages, an alternative approach to oral assessment in which the external examiner was not physically present at the time of assessment would make the introduction of aural assessment possible.

3.9 Recommendation

Consideration may be given to the introduction of an aural assessment developed centrally and provided to ensure that all four language skills are assessed.

3.10 Roles and responsibilities of key stakeholders

It is important that an examination system, as any other quality system, operates with clearly defined roles and responsibilities and that these contribute to the effective and efficient organisation of the system as a whole. In the case of the European Baccalaureate, the key operational roles are those of the teachers and Heads of Department, the examiners and the Inspectors. In this section we review each in turn.

3.10.1 Teachers and Heads of Department

Teachers in the European Baccalaureate system play a vital role in the effective delivery of the curriculum and in the conduct of significant volumes of internal assessment. The teachers' role as assessors is key to the ethos and functional operation of the system as a

whole. Our evaluation has noted the extent to which teachers are highly experienced in their classroom practice. However, the teaching body is made up of teachers in many different circumstances. Some are on secondment, while others are on local contracts (approximately 31.8%, see Section 3.7.2). Each year 10-20% of teachers are recruited new to the system. It is important that the roles assigned to teachers in terms of their assessment functions are supported with effective training and induction programmes. Well-designed activities to encourage the sharing of best practice across the network (see section 3.2.6) and to counter the degree of 'didactic individualism' (see section 3.3.2) are important in ensuring that all the advantages of internal assessment are delivered for the system.

Teachers' classroom assessment role is supplemented by two (and for some, three) additional assessment roles. Question paper production for final examinations (and oral assessments where they occur), first marking of those final examinations and assessment of oral examinations. Our evaluation has proposed revision of the first of these responsibilities so that teachers, while maintaining an involvement in setting question papers are not asked to design an entire question paper (see section 3.6.2.2). While we concur with the proposal that the construction of European Baccalaureate be outsourced to external experts, we consider that the continuation of teachers' involvement in question development adds value to the system overall.

In respect of teachers' involvement in the 'first marking' of final written examinations, we have expressed concern about this practice on a number of counts (see section 3.7.4), and suggested a revised approach whereby teachers' knowledge of their students is captured in the production of a class rank order of students rather than a system of marks (see section 3.7.6). There would appear to be acceptance of the principle that in double marking 'teachers give higher marks' and thus, an acknowledgement that an inconsistency of standard operates across the system. It would be useful to conduct a simple survey to demonstrate empirically the extent to which this statement is correct. If evidence supports the perception that teachers' marks are higher than those awarded by examiners, the potential exists for students' final marks to be overstated.

The involvement of teachers in oral assessments has the effect of significantly increasing the role that they play in contributing to European Baccalaureate outcomes. Teachers design and take a dominant role in the conduct of oral assessments. Again, we have suggested that this practice be reviewed and that the construction of oral assessments be revised to follow the same process as that of other final examinations (see section 3.8.3.2).

These proposals capture, we believe, the significant areas for review in the current roles and responsibilities of teachers in so far as they relate to the European Baccalaureate.

3.10.2 Examiners

The key responsibility of an examiner is to mark students' work fairly and consistently. Examiners contracted to provide external marking services in the European Baccalaureate examination are highly-experienced (see sections 3.3.4, 3.7.5). The majority of their experience tends to have been gained in examinations other than the European Baccalaureate, as it is frequently their status as national 'experts' that leads to their appointment by the Lead Inspector in a subject.

Examiners understand their role very well. What is less clear is the robustness of the quality assurance system that underpins their work. Their own marking is not routinely supported by a standardisation process at the beginning of the marking process. Procedures for reconciliation of discrepancies in double marking scores exist, but the concept of monitoring and review of marking as a process did not appear to be well-defined. Responsibility for the detection of 'aberrance' and marking review (other than the escalation to third marking with discrepant scores) appears to be weakly defined at a system level. The French term 'jury' captures this process very well (see section 3.2.7). A quality system needs assurance that outcomes are subjected to a process of review. In the case of the European Baccalaureate examination it would seem that the major instrument by which a review is conducted, the Annual Report of the Baccalaureate examination, takes place in November, some time after the issue of results. There is scope for the establishment of a clear monitoring and review process in European Baccalaureate marking with responsibility assigned to a panel of senior examiners or 'jury' able to review and report on the quality of individual examiner, teacher and subject marking, and outcomes across all subjects.

Such monitoring and review processes should be designed so that, if future pilots indicate that examining should be conducted away from a marking centre such as Brussels, a system is in place to monitor marking performance.

3.10.3 Inspectors

The role of the Inspector in the European Schools system is a crucial and very wide-ranging one. To quote the view of one Inspector, "we sometimes feel that we are inspecting our own work".

The Inspectors highlighted the differing amounts of time that each could spend on European Schools' work. For most the percentage ranged from 40–60% of workload. However, there are significant variations in the responsibilities of Inspectors. For some there was a national responsibility for quality assurance but very few teachers. On the other hand, the English Inspector had a responsibility to ensure the recruitment and quality of the work of 240 teachers in the system. In addition, some Inspectors are also given responsibility for a subject (for the quality of the syllabus, for its development and for its assessments).

Inspectors have a wide range of functions in terms of the curriculum and its delivery. The inspection of teaching quality, the development of protocols to ensure best practice and national compliance in relation to learning support, special needs and educational developments are all vital to their responsibility of quality assurance for teaching and learning. The report of the review carried out by the Inspectorate makes clear this commitment.[46]

This report also makes clear the Inspectors' role to "harmonise and co-ordinate curriculum development and to oversee the implementation process in the schools type I, II and III."[47] Inspectors' involvement in the quality of teaching and learning is key to quality in the system. Our evaluation has identified considerable variation in the process of syllabus revision and the implementation of curriculum development and Inspectors' commitment to a process of regular review and development is sound.

[46] Role of the Boards of Inspectors and Areas of Quality Assurance within their Fields of Competence, 2008.

[47] Section 2.1, ibid.

In addition to the wide range of roles associated with the curriculum delivery and development, Inspectors currently identify their responsibilities in relation to the "the quality and the coherence of the organisation of the European Baccalaureate in all types of schools" with a commitment that "every school will be inspected by secondary inspectors during the period of the written and oral exams. The Secondary Board of Inspectors decides about the nature of these inspections"[48].

Given the number of responsibilities that Inspectors must discharge to schools in terms of curriculum delivery, we consider that a review of the range of responsibilities associated with the European Baccalaureate might be considered.

In quality assurance systems as a whole there is a difference between monitoring the operation of a system and the direct inspection of all processes within it. While it is essential that during the examination period a process of checks and invigilation is carried out in all schools to secure the integrity of examinations there are a number of ways in which this might be done other than direct observation by Inspectors. Arrangements might be developed to require all schools to ensure that every examination is invigilated by a responsible person with no connection to any of the candidates and that an invigilators' report be provided to the Bac Unit. This is the practice in many school and university examinations. Further quality assurance may be provided by a system of security audits in which compliance with the regulations is assessed. The responsibility for ensuring that security auditors were in place would be one that the Board of Inspectors would wish, in all likelihood, to delegate. A number of metrics, such as the sampling rate of security audits, the criteria for review and the format of reports may be determined by the Board of Inspectors. The School Director, together with the Assistant Director responsible for the European Baccalaureate, would be responsible, as now, for ensuring correct operation of examination procedures.

An adjustment to the Inspectors' role such that responsibility moves from a direct role (physical overseeing each and every examination during the course of the session) to a quality assurance overseeing role involving monitoring of compliance would seem to deliver considerable advantage, as the time of the Inspectors can be prioritised towards the areas of curriculum and pedagogy which need to be addressed.

Such a redefinition of role from what might be called direct scrutiny (physical presence to observe a process) to quality assurance (the development and monitoring of a quality system) might also be applicable to involvement in the marking exercise (whether residential or dispersed). Currently, the Inspectors are present throughout the marking process whereas the unique value of their role is more closely linked to the range of activities defined in their report as "evaluate these results from a pedagogical point of view in order to develop the quality of the baccalaureate exams"[49] A more effective delivery of this responsibility might be designed to promote the analysis and evaluation of student outcomes in one session into the teaching and learning of the next.

3.10 Recommendation

In respect of the roles and responsibilities of stake holders we recommend:

[48] Section 3, ibid.

[49] Section 3, ibid.

- *review of teachers' responsibility for the construction of whole question papers*
- *review of teachers' responsibility for first marking (in favour of a rank ordering of candidates)*
- *establishment of a panel of senior examiners or 'jury' able to review and report on the quality of individual examiner, teacher and subject marking, outcomes across all subjects.*
- *Inspectors' direct scrutiny of the conduct of examinations in European Schools be delegated to invigilators and observers who are able to report on regulatory compliance to the Board of Inspectors*
- *Inspectors' direct scrutiny of the conduct of marking be delegated to the panel of senior examiners or 'jury'.*

4 Recognition of the European Baccalaureate in Higher Education

4.1 Overview

This chapter considers the nature of the recognition given to the European Baccalaureate and its effectiveness in respect of its predictive validity for Higher Education outcomes.

4.2 Research base

The research undertaken as part of this evaluation involved three main elements: general research into the recognition of European Baccalaureate within Europe; obtaining information from 10 higher education institutes (HEIs) in the UK and ten HEIs in France; on their admission policies and a range of contacts with other bodies. The outcomes of the Analysis of the Academic and Professional Careers of the European Schools' Graduates[50] survey have been published since this research work commenced and complement the findings of our study.

Although the European Baccalaureate is a European qualification, and European Baccalaureate students go on to universities across Europe and the world, the UK and France are the main destination countries for European Baccalaureate students. The other bodies contacted in the UK were UCAS, the national admissions service for entry to higher education, and UK NARIC, the national agency responsible for providing information on qualifications from outside the UK. In France, the equivalent agency, ENIC NARIC was contacted at CIEP (*Centre international d'études pédagogiques*). One of CIEP's 'main missions' in France being "to internationalise its education system".[51]

4.3 Research methods

Information was obtained in three ways: from the web, through submitting questions by e-mail and through telephone interviews and conversations. Each method had advantages and limitations.

The researchers, one in the UK and one in France, used the same recording sheet for information from HEIs, an example of which is attached as Appendix (4). This standardised the information required and by standardising recording, sought to facilitate comparison.

4.4 Summary of previously published information about equivalence

Three sources helped to inform the study and to give some background to it.

4.4.1 Of the Statute of the European Schools

Article 5 (2) of the Statute of the European School (1994)[52] provides that holders of the European Baccalaureate shall:

[50] Op.cit.

[51] Home page, About CIEP, English version. http://www.ciep.fr/en/index.php

[52] http://www.eursc.eu/fichiers/contenu_fichiers1/257/SW1-21994A0817en.doc

(a) enjoy, in the Member State of which they are nationals, all the benefits attaching to the possession of the diploma or certificate awarded at the end of the secondary school education in that country; and

(b) be entitled to seek admission to any university in the territory of any Member State on the same terms as nationals of that Member State with equivalent qualifications.

4.4.2 The European Schools and the European Baccalaureate

This information booklet, designed as guidance for universities and colleges was produced in the UK by the then DES, together with the European Schools. It is available from the successor to the DES, the Department for Children, Families and Schools (DCSF) website.[53]

The content is quite out of date, both materially and statistically, being based on data for the period 1998–2001. The booklet is undated but was possibly published in late 2002. At the time, it notes that approximately 400 candidates from the European Schools applied to UK HEIs each year.

In terms of equivalence with A level,[54] the following points are made:

- that progression is conditional in the European Baccalaureate, so pass rates are very high (96.5% 1999–2001)
- that its characteristic breadth means that candidates may be taking some subjects which they would have dropped if selecting A Levels. As such universities cannot reasonably set an offer of 6/10 in all 10 constituent subjects (given the pass mark of 60%).

Further, it states that UK universities may make three types of offer to EB candidates, specifying:

1. a final European Baccalaureate score above 60% (the pass mark)
2. a final European Baccalaureate score and marks in certain individual subjects
3. a pass in the European Baccalaureate and marks in certain individual subjects.

Further guidance is given on offers: “Institutions might consider that candidates should not be required to obtain 60% (or more) in all subjects; and that the number of individual subjects specified in an offer should not exceed three.”

There is also some broad guidance on English language proficiency for those who take English as a second or third language.

In terms of performance, there is a general comment on page 20, that, “degree results of EB students follow the national distribution of degree results by degree class. Drop out rates are similar too.” There is a small data set, for EB students who entered university in the UK in 1998, in the booklet’s Appendix 2. Although the data is described as “illustrative”, some

[53] Department for Education and Skills and Schola Europaea, *The European Schools and the European Baccalaureate: Guidance for Universities and Colleges* (undated).
http://www.dcsf.gov.uk/publications/pdf/dfes_baccalaureate_booklet.pdf

[54] Section 5, p15-19.

congruence between European Baccalaureate score and degree outcome is discernible in Table 4.1.

EB score (%)	Degree class				Dropped out
	1	2.1	2.2	3	
90–94.99	3	-	1	-	-
85–89.99	4	10	3	1	-
80–84.99	-	6	2	-	1
75–79.99	2	8	3	-	-
70–74.99	2	5	5	1	1
65–69.99	-	2	5	-	-
60–64.99	-	-	1	1	2

Table 4.1 European Baccalaureate score and class of degree awarded for 67 entrants to UK HEIs in 1998

4.4.2 Recommendation

That the Guidance Document provided for UK universities by DCSF be swiftly revised. Its information is now out of date. Its reference to equivalent standards in particular in respect of Science A Levels may no longer be appropriate.[55]

4.4.3 A study of the performance of European Baccalaureate students in Higher Education in the UK and Ireland[56]

As the literature about the European Baccalaureate is limited, this study is often referred to. However, whilst interesting, it appears less than robust in terms of design and pursuit and its findings have, therefore, to be treated with caution. The authors recognised, for example, that good students were over-represented in the survey.

The study aimed "to compare A-level grades and European Baccalaureate results, with a view to informing the selection process at British universities"[57]. The study proposed a level of equivalence between European Baccalaureate scores and A level grades[58], shown in Table 4.2. These equivalences were derived using the percentage of candidates achieving firsts and upper seconds, termed "good degrees", as the basis of comparison between the two qualifications. However, no critique of this approach was offered and no information was given on scores below 60%, a European Baccalaureate fail.

[55] See Section 3.4.3 of this report.

[56] Kelly, Daniel and Kelly, Alison, April 2006 http://www.yourschool.net/Survey/EB.pdf

[57] Ibid, p.1.

[58] Ibid, p. 5.

European Baccalaureate B score (%)	A Level grades
80 –	AAA
79–70	AAB, ABB
69–60 [sic]	BBB, BBC
60–64	CBB, CCB, CCC

Table 4.2 Equivalence of European Baccalaureate scores and A Level grades (Kelly and Kelly, 2006)

4.4.4 The academic and professional careers of graduates of the European Baccalaureate

The survey conducted by Van Dijk on the Academic and Professional Careers of Graduates of the European Schools provides significantly more comprehensive information on a data set of just over 3000 responses or about 8.3% of all graduates of the European Baccalaureate. This study provides a strong indication of the predictive validity of the European Baccalaureate. Using an A-E classification system to define outcome performance at a higher education institution and cross-referencing with students' European Baccalaureate scores produced the following table:

Extract: Academic and Professional Careers of Graduates of the European Baccalaureate.

Degree at college or university (classified by performance level, A highest, E lowest) →		**A**	**B**	**C**	**D**	**E**
Degree for European Baccalaureate ↓						
	60-64%	**38.7%**	**32.3%**	**17.7%**	**6.5%**	**4.8%**
	65-69%	**44.3%**	**30.7%**	**12.5%**	**2.3%**	**10.2%**
	70-74%	**29.1%**	**49.0%**	**16.6%**	**2.0%**	**3.3%**
	75-79%	**44.3%**	**43.7%**	**9.0%**	**2.4%**	**0.6%**
	80-84%	**50.7%**	**39.6%**	**7.5%**	**1.5%**	**0.7%**
	85-89%	**49.4%**	**40.3%**	**7.8%**	**1.3%**	**1.3%**
	90% or more	**72.7%**	**22.7%**	**4.5%**	**0.0%**	**0.0%**

Table 4.3: Cross relation between degree realised at secondary (European Baccalaureate) and tertiary (college or university) level

As the Van Dijk survey points out, about 95% of graduates who obtained their European Baccalaureate with a score of 90% or more obtained their university diploma with the highest (72.7%) or next highest (22.7%) scores possible. These levels of predictive validity are higher than for many national qualifications, for example, A Levels in the UK, demonstrating the

effectiveness of the European Schools educational system. Taking account of the fact that those who were just over the minimum for obtaining the European Baccalaureate might still go on to perform well at university, it would seem, as the authors of the Van Dijk survey report, "that the European Schools' system lays a good fundament for succeeding in tertiary education."[59]

4.4.4 Recommendation

That guidance documents for university admissions purposes and European Schools website information are revised using information supplied in the Van Dijk study.

4.5 Universities as a source of information about the European Baccalaureate

Some, but not all, universities in European countries other than the UK and France, give information about the European Baccalaureate on their websites. For example, the Université Libre de Bruxelles – Université d'Europe[60] states the position from the Statute (see 3.1) and is clear that no *demande d'équivalence* is necessary. By contrast, TU Delft, Delft University of Technology, Netherlands, recognises a list of qualifications, including the European Baccalaureate, IB, A Levels and the French *Bac S*, as equivalent to the Dutch pre-university VWO school diploma. It is specific about admission requirements for European Baccalaureate students to each of its Bachelor of Science degree courses.[61] For all programmes in Dutch, it specifies European Baccalaureate subjects as 'Dutch + Mathematics (5) + Physics". Two courses have Chemistry as a further requirement.

A small number of enquiries, in English, to universities in other EU member states, received no responses.

4.6 The European Schools websites as sources of information about the European Baccalaureate

The material available on the websites of the 14 European schools for those applying to university varies. For example, Luxembourg (I) has practical information for European Baccalaureate students who wish to apply to universities in France,[62] whereas Bruxelles (II) has a detailed table of equivalence,[63] reproduced in Table 4.3, but no information about its source or how it was derived. It is based on the points system which gives Grade A 120 points, B 100 points, C 80 points, D 60 points, E 40 points and F 0 points and ranges from an AAA equivalence of 84% through all combinations to a DDE profile equated with 60%.

The table employed by Brussels II is more detailed than the findings of Kelly and Kelly in Table 4.2, but offers different equivalences, for a different purpose. It is not known how widespread the use of the information in this table is within the network of European Schools.

---

[59] Op.cit. p43.

[60] http://www.ulb.ac.be/enseignements/inscriptions/bacEuro.html

[61] http://w76ww.tudelft.nl/live/pagina.jsp?id+b8abeca1-74f7-4c46-a59b-e247b18ed4ee&lang=en

[62] http://ww75w.euroschool.lu/orientation_france/pages/longues.htm

[63] http://www74.eeb2.be/site/index.php?id=296&L=4

4.6 Recommendation

That information on European Schools' websites should be standardised across all schools of the network.

4.7 Study Findings

4.7.1 Universities and Colleges Admissions Services (UCAS) in the UK

Given that the UK represents a significant destination for European Baccalaureates the evaluation considered the Information provided by UCAS, the UK central agency for overseeing university admissions. Information from UCAS may be divided into two types: that which is available to the public on its website, and that which UCAS makes available to HEIs, for example in its publications.

Visiting the website, the European Baccalaureate does not have a UCAS tariff and does not appear in the list of qualifications, 28 in total.[64] "Non-UK advisers", i.e. those from overseas, are simply referred to UK NARIC (see 5.2). However, UCAS does provide 2 pages of information about the European Baccalaureate in their publication *International Qualifications*. It is mainly a description of the qualification with a short paragraph on what is termed acceptability. Under this, admissions decision-makers are encouraged to remember the compulsory nature of much of the Diploma and that applicants may be of other nationalities, who have completed their studies in a language other than English. There is no statement of equivalence or guidance on comparability.

4.7.1 Recommendation

That revised text for the UCAS publication 'International Qualifications' be submitted drawing on information derived from the Van Dijk Report.

4.7.2 UK NARIC

The European Baccalaureate falls outside the remit for the UK NARIC as, through not being a national qualification, it does not fit the organisation's evaluation criteria. Access to the UK NARIC website is for members only. The website is established on the basis of "country files" using a world map and alphabetical list of countries, into which a number of qualifications, including the European Baccalaureate, do not fit. A manager at UK NARIC was keen to point out that this is not to suggest a criticism of the European Baccalaureate, nor that it cannot be used for entry to UK HEIs. Individuals who enquire are advised that it "meets admission requirements", but are advised to deal with each institution separately.

UK NARIC's member organisations can access information about what the European Baccalaureate is in one of a number of appendices. However the information is not found by the site's own Search facility. A manager at UK NARIC pointed out that such a description of the qualification is not a statement of equivalence, and that it is not possible for them to provide grade comparability. Instead, if required, they undertake to explain how the grading is done, in order for the member institution to come to its own understanding and conclusion.

[64] http://www.ucas.ac.uk/advisers/curriculumandquals/ucas_tariff/qualifications/

4.7.2 Recommendation

We recommend that a dialogue be established with UK NARIC to improve information and understanding.

4.7.3 Universities

The admissions departments of 10 universities were contacted. In alphabetical order the universities were: Bristol, Cambridge, Glasgow, Imperial (London), Nottingham, Oxford, Sheffield, Southampton, UCL and Warwick. All are members of the Russell Group, an association of 20 leading "research-intensive" universities in the UK. Of the 10 contacted, 9 are in England and 1 is in Scotland.

4.7.3.1 Recognition

All 10 universities recognise the European Baccalaureate for entry.

At Nottingham, although the European Baccalaureate is "welcomed", the European Baccalaureate did not appear in the list of 'Qualifications accepted' on the website.[65] When this was reported, the International Officer for EU Students took steps to have the list amended and the European Baccalaureate included.

4.7.3.1 Recommendation

We recommend that an awareness-raising campaign be conducted with key UK university International Admissions Offices.

4.7.3.2 Records and data

It was only possible to obtain data about the European Baccalaureate from 3 of the 10 universities. This was because most of their systems were not set up in such a manner that it was possible to identify applicants and students using the European Baccalaureate for entry. This response, from one Head of Admissions, typifies those received:

"Unfortunately, UCL does not record the names and numbers of students that enter UCL with European Baccalaureate qualification. For us to assist with your evaluation, it would mean having to search through all our student files and then, once students with European Baccalaureate qualifications had been identified, undertake investigations into the performance of these students. I regret that whilst this would be of interest to us, we cannot resource such an exercise at present and will not therefore be able to assist you."

A limited amount of data about applications was obtained from Bristol, Oxford and Warwick.

At Bristol in the 2008 cycle, of the 42,862 applications in total, 138 were from European Baccalaureate students. Of these, 5 received unconditional offers, 55 received conditional offers and 78 were unsuccessful. Of these, at time of writing, 11 were accepted for entrance in 08/09 and 09/10 (after a gap year). In the 2007 cycle there were 9 acceptances.

Admissions data for Oxford, published annually, is shown for 2002–7 in Table 4.4

[65] http://www.nottingham.ac.uk/ugstudy/other-qualifications.php

Year of entry	Total applicants	EB	French B*ac*	IB
2002	11,097	8	9	97
2003	11,793	9	8	99
2004	12,278	6	8	114
2005	12,496	6	20	135
2006	12,614	9	11	135
2007	13,639	7	18	139

Table 4.4 University of Oxford: Conditional offers given to Pre-qualification applicants

Of these 6-9 offers, between 0 and 2 did not take up their places. Up until 2006, all applicants were from Culham, Bruxelles and Luxembourg. In 2007, there were applicants from Munich for the first time, coinciding with a campaign by the Undergraduate Admissions Office to recruit more students from continental Europe. Oxford is a collegiate university. At the College level, 2003–2008 records show that 20 of the 30 colleges made offers to candidates with the European Baccalaureate, 14 of them making more than one offer, and 2 colleges (St Hilda's and St Peter's) each making more than 5 offers.

In the 2008 admissions cycle, Warwick received 56 applications from candidates presenting the European Baccalaureate, of whom 37 were made offers. In 2007, "approximately 30%" of European Baccalaureate offer-holders went on to have places confirmed at Warwick for taking undergraduate degrees.

4.7.3.3 Offers for entry and comparability with other qualifications

Findings from the 10 UK universities contacted are diverse. The status of the European Baccalaureate as an entry qualification varies markedly in terms of the information about it and its inclusion in admissions materials and course entry details.

It should be noted that the AS and A Level system in England and Wales has changed with the introduction of new specifications for teaching from September 2008, with first AS award in June 2009 and first A2 award in June 2010. Changes include the introduction of a new A* grade at A2, to identify higher-achievers, in part to assist admissions to universities. Findings in this study relate to admissions cycles up to and including 2009.

The overall guidelines from Oxford and Cambridge about likely kinds of offers that are made, or achievement necessary, are given in Table 4.5.

	European Baccalaureate	French *bac*	A Levels
Oxford	Average of 85% or above with scores between 8 and 9 in specified subjects.	An average score of at least 15.	AAA (not in General Studies)
Cambridge	80-85% overall, with 90% in the subjects closest to those to be studied.	16 or 17 overall and usually asked for 16 or 17 in specific subjects.	AAA (some with AEA*)

Table 4.5 Outline requirements for entry to the Universities of Oxford and Cambridge

The clearest statement of equivalence is information for the staff of Nottingham University, which was made available to this study. The "suggested equivalencies" given in Table 4.6, have been in place for a number of years, are of unknown origins and compare well with Table 4.3. They are regarded by Admissions staff as good, but perhaps in need of updating.

EB (%)	A Level grades
85	AAA
80	AAB
75	ABB, BBB
70	BBC, BCC

Table 4.6 Suggested equivalencies, University of Nottingham (for staff use)

Evidence from other universities is more complex, less unified and at times apparently contradictory between courses, departments, schools or faculties. Two courses, Chemistry and History, available at most UK universities, are shown in Table 7 as examples.

University	Degree course	EB score (%)	French *bac*	A Levels
Bristol	Chemistry F100	70-73%	Accepted, but not listed.	ABB - BBC
	History V100	80-85%	Accepted, but not listed.	AAA - AAB
Glasgow	Chemistry (all sciences)	65% including 3 science subjects	Not listed.	BBC to BCC
	History (all arts and social sciences)	75%	Not listed.	ABB
UCL	Chemistry F100	80%	"with good grades"	AAB to BBB
	History V100	85%	15	AAA including History

Table 4.7 Comparison of entry requirements for Chemistry and History at 3 universities in the UK

4.7.3.4 Language proficiency in English

The 10 UK universities vary both in what they seek from European Baccalaureate students for language proficiency in English and how they express their requirements. Outline details of these are given in Table 4.8.

University	Statement on language proficiency for entry to undergraduate degrees using European Baccalaureate
Bristol	All whose first language is not English need to demonstrate proficiency in an approved language test. 'European Baccalaureate with a minimum 7.5 in English as the First Foreign Language L2 or First Foreign Language Advanced – L2A'
Cambridge	'English as a compulsory or option subject must have 8.5 or 85% minimum' (Admissions Forum agreed list, 27 January 2006)
Glasgow	'Candidates whose dominant language is not English must provide proof of proficiency in the English language, achieved within 2 years of application.'
Imperial	All entrants "must have full command of the English language". European Baccalaureate minimum requirement 6.5 in English.
Nottingham	Non-native speakers require certification, such as IELTS or TOEFL, but this is waived for European Baccalaureate candidates who have English as their First or Second Foreign Language and achieve a 6 or a 7 (viewed as equivalent to IELTS 6, 6.5 and 7).
Oxford	All non-native speakers of English need to demonstrate language proficiency. European Baccalaureate score of 70% in English.
Sheffield	"If English is the First Language studied and passed at 60% or higher, no further certification is required."
Southampton	'7.5 English as the First Foreign Language L2 or First Foreign Language Advanced L2A'
UCL	'8 in English as either an Obligatory Subject (Language I or II) or as a full Option (Language III or IV).'
Warwick	For students taking English as a subject, depending on the course, "offers will stipulate English language achievement scores currently ranging from 5-7." No additional language test is needed.

Table 4.8 English language proficiency requirements for European Baccalaureate applicants to 10 UK universities

4.7.3.5 Performance of European Baccalaureate students on degree courses

Performance data linked to entry qualification appears not to be kept routinely by 9 of the 10 universities contacted and none of the 9 had conducted a recent systematic review of performance to be able to respond. As such the few comments about performance received were broad, such as "we have never had any problems in the past with European Baccalaureate students" (Nottingham).

The exception to this was Oxford. All 15 students who entered using European Baccalaureate and completed their degrees in 2006, 2007 and 2008 attained the class 2.1. This was for subjects ranging from Physics to History. Only one student (Modern Languages) withdrew during that period.

4.7.3.5 Recommendation

We recommend that European universities are informed of a summary of the Van Dijk outcomes showing high levels of predictive validity.

4.7.3.6 Other observations from UK universities

It is notable that all admissions contacts in the 10 universities were positive about the European Baccalaureate as an entry qualification. Comments included, "the European Baccalaureate is well-regarded," (Bristol); "all our Schools welcome European Baccalaureate students for the breadth of their study", (Nottingham); and, the European Baccalaureate is "very welcome for entry to all our courses" (Page 6, *Undergraduate Admissions Policy,* University of Sheffield).

Two universities, Oxford and Cambridge, responded at greater length.

"The European Baccalaureate appears on the face of it [i.e. on the basis of the performance data at 5.3.5] to be a suitable preparation for Oxford – it is encouraging to see successful students in the sciences, humanities and social sciences, as this would indicate that the European Baccalaureate can equip students equally well across the board. We would be keen to see more applicants of the calibre we are admitting apply." Mike Nicholson, Director, Undergraduate Admissions, University of Oxford, (e-mail 30 July 08)

The University of Cambridge Admissions Office (CAO), lists on its website Admission Tutors and others who have "particular expertise". Dr Paul Russell (Pembroke) is shown for the European Baccalaureate. Although many of the Colleges' admissions tutors are familiar with the European Baccalaureate and do not need his help, he advises those who are unsure about it. The following points emerged in discussion with him:

- the candidature is highly variable in character, coming from diverse backgrounds, different countries and schools which themselves differ;
- as such, it is hard to generalise about the European Baccalaureate, and with small numbers involved (which tend to make individuals memorable), each candidate needs to be taken on their own merits;
- some Cambridge offers appear high, compared to A Levels, for example requiring 8 or 8.5 in specific subjects, sometimes as high as 9.5 (exceptional) in the sciences;

- there may be a tendency observable for Cambridge to attract the very high fliers from European Baccalaureate schools, but maybe not enough of those European Baccalaureate students who will go on to achieve 2.1s;
- the candidates' breadth of preparation, workload and international-ness make European Baccalaureate candidates different and out-of-the-ordinary;
- and the nature of the assessment they have undergone (such as 3-hour examinations, extended 'Thought' type essays, orals, etc.) prepares them well for university, but may not always be well understood for admission.

4.7.3.6 Recommendation

We recommend that a list of officers with assigned responsibility for the European Baccalaureate be compiled for use by European Schools.

4.8 Study findings – France

4.8.1 *Centre ENIC NARIC France* at CIEP, Sèvres

The official response from ENIC NARIC was straightforward and clear.

"The qualification is recognized by all Members of the European Union, it gives access without any restriction to all the European universities.

Therefore, in France, a holder of a "European Baccalaureate" doesn't have to ask for an equivalence process and can enter any university like the holders of the French "Baccalauréat". (e-mail 29 August 08)

In a later conversation with one of the staff members, the divergence of universities in France from this stated position was recognised, given the relative independence of institutions and that some may not approach ENIC NARIC for information and assistance.

4.8.1 Recommendation

We recommend that a dialogue be established with ENIC NARIC to improve information and understanding.

4.8.2 Universities and Grandes Ecoles

10 selected HEIs were contacted. As France has a dual system, these can be divided into *Grandes Ecoles* and universities. There were 4 *Grandes Ecoles* contacted: *Ecole Polytechnique*, *Institut d'Etudes Politiques de Paris* (also known as *Sciences Po*) and the *Ecole Normale Supérieure* (Paris and Lyon). The universities were: Grenoble 1, Montpellier 2, Paris 4, Paris 6, Paris 11 and Strasbourg 1; making 6 in total.

Contacting the institutions was hampered initially by the summer vacation in France with the closure of offices from mid-July until mid-August or later.

4.8.2.1 Grandes Ecoles

The prestigious *Grandes Ecoles* are entered through a two-track system and are separate from the rest of the university system. *Voie 1* (Track 1) is for those who have attended post-secondary *classes préparatoires* (CPGE). *Voie 2* (Track 2) is for the "most promising" students who do not apply to these institutions until they are in the final year of their undergraduate studies (in the Humanities) or in their second or third year (in the Sciences). It is therefore a second stage admission and so is of less immediate relevance to this study than *Voie 1*.

Those who took the European Baccalaureate can apply for classes préparatoires but there is no immediate information for those who do not hold the French bac on www.admission-postbac.org, the admissions website. These classes last for two, sometimes three, years and prepare students for a highly competitive entrance examination at the national level, the concours.

The *Ecole Polytechnique* replied that it is open to international students. It emphasised the need for a good knowledge of French, although 'good' remained undefined. It reported that each year it admits up to 400 French students and 35 international students through *Voie 1* (Track 1) from *classes préparatoires* after secondary school, and up to 10 French students and 75 international students through *Voie 2* (Track 2), after or part way through, studies at university. No data about entrants with a European Baccalaureate background were available.

At *Sciences Po* in Paris, the contact replied that international students can enter the *premier cycle* (first cycle), which consists of three multidisciplinary foundation years, straight from secondary school "if they have obtained their baccalaureate or its equivalent". They may also enter after one year in higher education. International students prepare an electronic portfolio or dossier for entry and have an interview. *Sciences Po* also offers *le Programme International*, an international programme, which can be entered after two years of university study outside France.

These entrants study alongside those who have come through the first cycle. The programme is delivered in French, in English or in a combination of the two languages. Whilst it is easy to see that high-achieving European Baccalaureate candidates could be well-suited to this, no data was forthcoming for whom *Sciences Po* had admitted.

The *Ecole Normale Supérieure*[66] replied that, "There are no requisite programs of study for this application process. Through a series of tests, the selection committee evaluates each candidate's capacity for analysis, conceptualization and synthesis. Also taken into account are the candidate's scientific and literary knowledge, as well as his/her intellectual curiosity and the relevance of his/her chosen project."

The responses of the *Grandes Ecoles* emphasise the importance of liaison with *classes préparatoires*. Consideration might be given to a more extensive survey in this respect.

4.8.2.2 The universities

It was notable that most institutions seemed not to know what to with the request for information and passed the researcher from one office to another, for example from *relations*

[66] See website www.ens.fr

internationales to *scolarité* and on to the *bureau des étudiants étrangers*. This itself suggested a lack of familiarity with the European Baccalaureate as a qualification.

At *Université Joseph Fourier,* Grenoble 1, the European Baccalaureate is recognised for entry. Very few applications have been received and no data were kept. No specific marks are required; the response being that *"L'inscription se fait de droit"*, i.e. enrolment is by right. This broad answer did not address the issue of particular requirements for entry to courses such as medicine.

At Montpellier II, *the European Baccalaureate* is recognised. Some courses have specific requirements and attention is paid to the major subjects, for example mathematics and sciences for a mathematics course, and to French. There were no applications from European Baccalaureate students in the cycle of admissions for 2008, but the university usually receives one or two per year. Montpellier II acknowledges the "*équivalence de droit entre le Bac européen et le bac français*" as previously outlined (3.1).

At Paris 6, *Université Pierre et Marie Curie*, the contact responded that, logically, the European Baccalaureate should be recognised for entry, but that their information was that no candidates had ever applied.

The contact at Paris 11 had no idea what the European Baccalaureate was and said that it was not recognised for entry to the university. Foreign students are required to take an entrance examination. Unless there are special circumstances which are not immediately apparent, this appears to contravene the Statute (3.1) which is simply phrased in terms of Member States of the European Union.

The initial contact at Strasbourg 1 was not aware of the existence of the European Baccalaureate and further enquiries elicited no other response.

4.8.2.2 Recommendation

We propose that an awareness-raising campaign be conducted with key French institutions.

4.8.3 Conclusions about recognition within France

The recognition of the European Baccalaureate amongst HEIs in France appears to be variable and contradictory. Replies to the survey and knowledge of the European Baccalaureate seemed to depend very much on the institution's or respondent's experience. European Baccalaureate applicants could benefit from greater awareness of the European Baccalaureate as a qualification, especially if seeking to enter an institution unaccustomed to receiving such applications, which has not sought advice from ENIC NARIC.

No statistics were kept by any of the institutions contacted, nor were records of performance for those who had entered using the European Baccalaureate. As such there was no point of comparison with UK HEIs.

Whilst in theory the European Baccalaureate should be recognised (3.1, 6.1), in practice the universities apply different procedures. Some grant "automatic equivalence", whilst others select using the marks obtained. In most cases non-francophone applicants have to take a French language test for entry to ensure genuine facility with French, both written and oral.

4.9 EB and admissions deadlines in EU universities

Students receive 2 reports with provisional results: one at the end of February (20 February this year for ES Karlsruhe) with the results of the partial exams of January and one at the end of May (26 May this year for ES of Karlsruhe) with the global preliminary results. This means that although European Baccalaureate students may not necessarily have their final Diplomas when they make their applications, they do have a considerable amount of information for universities to consider in making their decisions. The Diplomas are issued in early July - in 2006 July 7 and 8, in 2007 July 6 and 7 and in 2008 July 4 and 5.

The information about admission deadlines has focused on the countries with European Baccalaureate schools and some of the other EU countries that students may be more likely to be applying to. Admissions procedures throughout Europe vary enormously. The academic year starts as early as August in some countries and the end of October in others. In some countries applications are made before the school leaving examinations have been taken whereas in others the application is made post qualification. Even within one country there can be a surprising amount of variation and there are also differences depending on the faculty. It is common for courses with a high degree of competition, such as Medicine and Law, to require applications to be submitted earlier than for other courses.

The information provided is based on desk research through searching websites, and by contacting key organisations by email or phone. The information provided has come from a wide range of sources including the cultural section of embassies, ENIC/NARICs, universities, CIE regional representatives, CIE schools, and so on. Because of the variation and range of sources the information provided should be assumed to be representative but not definitive.

The European Baccalaureate examinations are over and the results released by early July. Therefore any application procedure that requires final grades before this may be problematic.

There appear to be three main scenarios for international students:

1. Applications are made before the examinations have been taken and the results are verified at a later date. This would include France, Italy, Luxembourg, Spain and the UK. In the UK applications are based on the preliminary results which include coursework and January examination results as well as additional information such as references and supporting statements. Universities make conditional offers which are confirmed after the final results are known. In Spain the Ministry of Education accept predicted grades from international students whereas local students sit the ‘selectividad’ examinations. Each country follows a similar pattern with local variations and may require additional documentation depending on the national admissions system.

2. Applications are made after the examination results are known but before the final deadline for application. This includes Germany where the deadline is July 15 and Belgium.

3. Applications are made after the examination results are known but the final deadline for application is the beginning of July. This includes the Netherlands. This category could be problematic. However, in some cases international students are allowed some leeway on this e.g. Norway and Switzerland, or if a higher fee is paid e.g. Austria.

In some cases the timing is quite tight and dependent on a degree of flexibility from the universities.

The situation becomes increasingly complex as more countries join the EU. Signatories to the Bologna process are increasingly bringing their processes into line in terms of start dates and so on but there is, nevertheless, a bewildering amount of variation in admissions processes. It is important that the school counsellors are familiar with the admissions procedures and deadlines for the specific countries their students are applying for.

4.9 Recommendation

Teachers should be advised to check carefully and ensure the applications are sent in good time so that arrangements for late verification can be made where required.

Given the extent of variation in procedures between and within countries, and indeed in some cases within individual universities, and given the lack of a 'clearing house' for information on this variation, the best resource would be the network of European Baccalaureate schools themselves. The systematic sharing of information regarding procedures in the systems to which their students apply would go a long way towards building up a complete picture for all countries, as a resource on which each school can draw. Such a database would need to be updated on a regular basis.

OSGES might also produce a briefing document or database to assist teachers in providing guidance to the students.

5 The European Baccalaureate and Expansion

5.1 Overview

In this section we review the opportunities for growth that are presented to the European Baccalaureate.

5.2 The integrity of the European Baccalaureate models for adoption

The European Baccalaureate is a 'total curriculum framework', it is not a suite of discrete qualification components which match a common set of criteria such as the A Level system in the UK, or Advanced Placement awards in the US. The fundamental issue in terms of greater participation is what model of the European Baccalaureate is to be considered for expansion. Three models may, different ways, be considered to retain the integrity and ethos of the European Baccalaureate.

5.3 Option Appraisal – Three models of adoption

Model A: Adoption of the European Schools Curriculum, in its entirety.

Model B: Adoption of the European Baccalaureate for students in years 6 and 7 alone.

Model C: Adoption of the 'European Baccalaureate' core curriculum (L1 and L2, History/Geography or Biology in L2) and mutual recognition of qualifications in other subjects.

5.4 Consideration of Model A

Model A continues the 'total curriculum framework' approach, which requires a wholesale commitment by a school. It is essentially 'all or nothing'. The adoption of the European Schools' curriculum is thus likely to be experienced as a 'paradigm shift' in educational policy. It requires a substantial incentive to effect the shift, and/or substantial catalyst. One such a catalyst has been provided by the closure of the JET agency in Oxfordshire and the proposal to re-establish the European School at Culham.

Given the topicality of discussion with respect to Culham it might be useful to take it as a Model A case study, albeit from the established footprint of a Type I school.

The April 2007 decision of the Board of Governors to close Culham as a Type I School provided a strong catalyst for the consideration of alternatives and current UK government policy offers a mechanism by which a Type III institution at Culham could be created.

The establishment of a system of Academies in the UK potentially provides a context in which the school would be established and this option is under current consideration. A European Academy would be able to operate within the state sector and admission to the school would remain free for parents. The proposed Academy would promote specialisation within the curriculum (in the case of Culham, Science and Languages). The Department for Children, Schools and Families (DCSF) in the UK has conducted a review both of transition arrangements to 2017 and the proposed Academy structure. There is strong support for the culture and ethos which the school would bring to UK educational policy to provide a flagship

for language teaching as a regional and national centre of excellence. A consultancy report carried out as part of this initiative[67] emphasises this:

"Alongside its role as a provider of education to 900 students in the school, the Academy will be tasked with promoting European languages teaching and learning across the UK" (para 67) and again, "The UK European Academy has the potential to lead the renaissance of language teaching in the UK state sector drawing on the expertise and experience of the European school staff and on the curriculum offered in the European Schools" (para 68).

Such statements make clear the potential which expansion into Type III schools offers the European Schools across the Member states.

However, against this background of strong commitment to the mission and ethos of the European schools' curriculum, it is apparent that the need for compliance with other dimensions of UK educational policy will result in the type of school envisaged at Culham being very different to other established European Schools.

A wide offer at 14-19 is assumed to be both essential and desirable and other qualifications will need to be offered alongside the European Baccalaureate. These could include GCSE, A levels and BTEC as well as language specific qualifications (para 71).[68]

An assumption that the European Baccalaureate may be accessed by "perhaps 60-70% of the new intake ranges" has a number of inevitable consequences for the curriculum:

- Likely adoption of external assessment at 16 by all (GCSE/GCSE/BTEC)
- The availability of a wider range of subjects in other curricular programmes than those offered by the European Baccalaureate is likely to produce a situation in which students seek to blend provision.
- The presence of single subject certification such as A level alongside the European Baccalaureate will require either additional teaching resources and small class sizes or co-teaching for alternative examination systems. A tension may be created by the triangulation of standards between A Level, European Baccalaureate assessment at Culham and European Baccalaureate assessment elsewhere.
- There will be a need to reconcile issues of curriculum coherence, plus
- L1 sections in German and French (Dutch/Italian) for which no alternative curriculum would be offered.

The European Schools' curriculum is a specialist and demanding one. Some students inevitably drop out, while some need to repeat a year. While repeat rates for the European Baccalaureate (16%)[69] are consistent with those in the EU as a whole, it is possible to predict that a significant number of students will find the level of Baccalaureate Mathematics (14.1% of all repeaters)[70] and Science (12.5%)[71] too demanding. The presence of alternative

[67] European School, Culham, Transformation to English National System as an Academy, Report by UK Government, Paul Doherty, DCSF, October 2008.

[68] Ibid.

[69] Academic and Professional Careers of students in European Schools.

[70] Ibid.

[71] Ibid.

curriculum programmes (GCSE, BTEC, A Levels) makes this a manageable situation for those in the English Language section at least if not for those in the proposed German or French sections.

Furthermore and inevitably, the multilingual nature of the environment would also be different; mother tongue teachers would not have the same cultural mix on which to draw; the context of the school will be very different in terms of its defining ethos. The need to offer other curriculum alongside that of the European Schools' and the presence of monolingual students will inevitably have an impact on the language culture of the school – and on its language mission.

"Everyday interaction in the playground, the corridors and the recreation rooms enhances the acquisition of the languages and using them is not only vital but natural."[72]

Culham as a UK European Academy would become very different to other schools currently within the network. It may be supposed that its emphasis on multi-lingualism and on the study of separate Sciences, together with the fact that admission would continue to be free will make the institution very popular with parents. It is uncertain whether there would be sufficient demand for full language sections other than English.

The range of issues which proposals for Culham present, give some indication of expansion based on 'whole school' adoption.

5.4 Recommendation

Working group proposals for Culham should give consideration to the range of advantages and disadvantages identified in Model A operation to ensure that detailed policy discussion can take place when further expansion opportunities arise.

5.5 Consideration of Model B

This option regards the European Baccalaureate as a year 6 and year 7 curriculum and assessment which schools that have not necessarily followed the European Schools' curriculum could be franchised to adopt.

This model is similar to that offered by the IBO in which a school can adopt the IB Diploma programme regardless of whether it has taken the earlier Primary Years Programme and Middle Years Programme. In this option, the European Baccalaureate Examination Board would need to establish a formal approval process by which schools would be accredited to offer the programme. This approach is that taken by Working Group 1: Accreditation of Schools in their April 2008 Report to the Board of Governors of the European Schools.

Appropriate quality metrics would be required to inform the approval process and extensive teacher training would be required to ensure that curriculum delivery and internal assessment

[72] Analysis of the Academic and Professional Careers of Students in European Schools.

was in line with the European Baccalaureate standard.[73] It is common for the IBO to require schools who are adopting the programme to commit to a specified volume of training.

In this option, schools would be accredited to offer the European Baccalaureate in Year 6 and 7. The number of language sections which were offered would depend on the Board of Governors' views about whether a single language section and a restricted (vehicular languages only) range of L2/L3/L4 choices would substantially change the ethos of the school. The recognition of the importance of multilingualism and the European Schools' acknowledged expertise in this area has the potential to offer significant opportunities for adoption of the European Baccalaureate across Europe.

Certainly, take up for the Model B option has a wider range of opportunities than for the adoption of the entire European Schools curriculum as described in Model A. However, the move to operating in L2 is currently introduced in Year 3 of the secondary curriculum. Its introduction at a late stage may substantially impact on standards. There are other implications. Currently, it is possible for some policies and procedures within the European Schools to be implicit because of the strength of the network as a whole. In order for one part of the curriculum, the European Baccalaureate programme, to operate independently of the mission and ethos of the schools as a whole, much more definition must be given to the overall programme aims and objectives; to the design of individual syllabuses in order to make explicit aspects, such as assessment objectives and their relative weightings. In particular, a training programme for teachers before they embark on the course and before they commence assessment would be required.

The funding model for such adoption is to charge an initial accreditation fee, and then to develop a services structure such that training and individual examination fees are invoiced separately.

5.5 Recommendation

Detailed consideration should be given to the range of advantages and disadvantages identified in Model B operation. Further discussion with agencies charged to develop EU multilingualism strategies might be considered.

5.6 Consideration of Model C

If the European Baccalaureate Examination Board views itself as an awarding body it has scope to redefine the components of the assessment for which it will make an award of the European Baccalaureate. Established exchange mechanisms that exist to secure the recognition of the European Baccalaureate can be developed so that the European Baccalaureate Examination Board recognises a range of appropriate national matriculation qualifications as satisfactorily meeting Baccalaureate requirements.

Students would thus obtain a national school-leaving qualification with the potential for additional endorsement via the European Baccalaureate.

[73] The Report of Working Group 1: Accreditation of Schools to the Board of Governors of the European Schools 15/16 April 2008 makes no reference to the training and approval of teachers as part of the Accreditation procedure.

An approval and accreditation process would need to be established, particularly to establish a calibration between the grades of the Baccalaureate and levels of achievement in other systems.

The adoption of elements of the European Baccalaureate curriculum alongside the matriculation qualifications of other educational systems provides an effective response to current goal-setting in connection with multi-lingualism and EU language policy. The range of language syllabuses which have been developed for the European Baccalaureate together with its expertise in assessing students in L2, L3 and L4 places the European Schools in a very strong position to lead on this development in Europe. An example to demonstrate the way in which this might work is provided by the Option Internationale of the French Baccalaureate. Principally designed for the children of different nationalities resident in France and taking the French Baccalaureate, it has many similarities with the emphasis on Language 2 and the approach to History/Geography adopted within the European Baccalaureate. Literature is also offered.

Organisation and administration of the annual examinations for the international option of the French Baccalaureate is delegated to the CIEP, under an agreement with the Delegation for International Relations and Cooperation (of the French Ministry of Education). The OIB thus becomes one of the French 'general' baccalaureates, with an additional international option. The Language sections leading to the OIB are available in Danish, Dutch, English, German, Italian, Portuguese, Spanish and Swedish, as shown in the table below.

SECTION	Passed in 2005	Passed in 2004	Passed in 2003	Passed in 2002	Passed in 2001	Passed in 2000
German	46	55	83	61	81	78
American	295	281	199	208	194	157
Arabic	8	7	3			
English	382	351	281	272	241	219
Danish	4	2	2	0	2	6
Spanish	205	223	197	207	168	164
Italian	92	80	78	71	65	67
Japanese	14	17	3			
Dutch	15	16	18	9	11	10
Polish	20	9	8			
Portuguese	39	37	17	27	37	13
Swedish	11	12	11	2	13	12
Total	**1 131**	**1 090**	**900**	**857**	**812**	**726**

Table 5.1 Participation rates across Language Sections 2000-2005 in the OIB

Responsibility for delivery of the examinations of the English Section is delegated to University of Cambridge International Examinations. Other national examination boards take responsibility for other subjects within the OIB.

In this model a school wishing to offer students the European Baccalaureate to its students would itself need to be accredited to do so. It would also then need the syllabuses which it was offering (from a national/international provider) to be accredited as meeting the quality standards of the European Baccalaureate. The list of approved syllabuses would be maintained centrally. Individual regulations about the non-examined component of the curriculum would need to be established. It would also be necessary to ensure that the prevailing ethos and character of the European Schools/curriculum is secured through a defined process of training.

5.6 Recommendation

Detailed consideration should be given to the range of advantages and disadvantages identified in Model C operation. Further consideration of the model presented by the inclusion of the Option Internationale within the French Baccalaureate would be of value – in which a European Option might be provided within the matriculation qualifications of other countries.

5.7 Alignment with Common European Framework of Reference for Languages

In the context of all expansion models it might be considered appropriate to develop more explicit links with the Common European Framework of Reference for Languages.[74] The core option of the European Baccalaureate particularly requires L2 linkage (and potentially L1 and L3/L4 linkages) with the CEFR. Our own curriculum mapping, together with the mapping conducted by the Van Dijk report has indicated relationships that exist in terms of Levels C1 and C2 for L2 and L1 respectively and B2 for L3. These are initial findings on our behalf and, as mentioned in Chapter 2 of this evaluation, the design of syllabus documents makes expected outcome levels rather difficult to determine.

The CEFR has exerted a considerable influence since its publication in 2001 on language teaching in Europe, at least at the policy level. Moreover, the adoption by the European Union of the CEFR as an educational performance indicator can itself be expected to strengthen its influence on language policies. The outcomes-based approach, a by-product of can-do descriptor scales, enables a clear 'read-across' with other systems. The emphasis on language use in a socio-cultural context fits very well with the holistic approach developed by the European Schools. The CEFR scales can act as a rich resource in syllabus and assessment design. A linkage of European Baccalaureate Language Syllabuses to the CEFR would provide a strong alignment with EU multilingualism strategies designed to raise performance levels above the almost 50% of European citizens who are monolingual.[75]

[74] See Chapter 2 of this report.

[75] Eurobarometer survey in 2006 cited in Commissioner Orban's speech: 'A new strategy for multilingualism: a strategy for all EU citizens', 6 October 2008, Brussels.

5.7 Recommendation

In order to underpin further expansion of the European Baccalaureate, we recommend that a more clearly defined relationship with the Common European Framework of Languages is established.

5.8 Impediments to adoption of the European Baccalaureate

In considering the models presented in this section it is also necessary to consider impediments to adoption.

5.8.1 Tension with National Systems

The assessment model of the European Baccalaureate is very distinctive. Not only does expansion place considerable stress on the European Baccalaureate administration (to provide an adequate number of external examinations), it also places considerable demand on school staff, who – depending on the characteristics of national assessment in which a specific school is located – may be unfamiliar/uneasy with the role of teacher assessment and the moderation and control arrangements associated with the European Baccalaureate. In some national settings, the emphasis on external assessment in the national system and in society may be at odds with the balance of assessment in the European Baccalaureate, thus generating tensions between the school and the dominant system/social culture.

5.8.2 The effect of coherent and distinctive ethos

The overt ethos of the qualification is very strong. Throughout the evaluation, teachers and administrators have emphasised the 'family' nature of the group of schools, the network of external examiners, and the administration. It has the appearance of a closely-knit community, sharing a strong sense of specific, and in many aspects unique, educational purpose. If the European Baccalaureate community emphasises that new schools can only operate the qualification in line with the approach described in Model A, then this is likely to be an impediment to present expansion. Not only are ethos issues difficult to express in formal participation/qualifying criteria, the sense of educational purpose frequently is a characteristic of strong management in schools[76] – it is considered in research to be (i) fundamental to the performance of the school as a social entity; and (ii) is seen as a 'territory' issue where head teachers and senior managers can make an impact on the school's performance.

In strong 'performativity' cultures,[77] school ethos is strongly associated with performance defined in terms of qualifications outcomes[78]. It is important that expansion is monitored to ensure a good fit between culture and ethos in new schools adopting the European Baccalaureate.

[76] Schmidt W and Prawat R, 2006, Curriculum Coherence and National Control of Education, Issue or Non-Issue?, Journal of Curriculum Studies, Vol 38, Number 6, 641-658.

[77] Ibid.

[78] Ibid.

5.8.3 Contractual issues

Practical aspects of the administration of the European Baccalaureate which are seen by the European Baccalaureate to relate to the ethos and key characteristics of the qualification, such as short term contracts, may be problematic in certain national settings (employment law, forms of employment etc) and against the ethos of some schools. An audit of these practical aspects in relation to support for the ethos of the European Baccalaureate might be undertaken by current European Baccalaureate stakeholders, in order to establish:

- which elements of ethos are fundamental to the European Baccalaureate and cannot be foregone
- which practical aspects of the European Baccalaureate support which elements of the ethos

This will allow management of the expansion – allowing some movement in the characteristics of the qualification in order to maximise the uptake in new schools, but without unduly affecting the ethos and purpose of the qualification.

5.8.4 The importance of value and recognition

Currency in respect of progression is critical. Understanding and recognition of non-typical qualifications is uneven in some systems and unusual qualifications may be subject to 'unfair' comparison with domestic and/or dominant qualifications. Unless recognition is clearly stable, fair and widespread – for all the major progression routes from the European Baccalaureate – it will operate as an impediment to expansion. The precise pattern of recognition of the European Baccalaureate is examined in Chapter 4 of this report.

5.8.5 Resourcing issues

The availability of appropriately trained, experienced and 'sympathetic' staff is a crucial matter. This issue is critical to expansion in two respects (i) the availability in specific national settings of appropriate numbers; (ii) the degree of reconfiguration of staffing which is required by a new school wishing to participate in the European Baccalaureate. Outside the European Baccalaureate, while the adoption of one new subject specification can include a requirement for new staffing (e.g. skills-oriented teachers in science), the European Baccalaureate is likely to require a substantial re-profiling of staff in schools. There may be considerable impediments to effecting this quickly and easily (e.g. national labour market agreements etc) and contributes further to a sizeable endeavour for schools wishing to adopt the European Baccalaureate for the first time.

5.8.6 Curriculum coherence

Key trans-national analysis asserts that 'curriculum coherence' is a characteristic of high performing educational systems.[79] This analysis is compelling, and draws on TIMSS and PISA data. Conversely, those systems not possessing curriculum coherence perform relatively poorly. 'Curriculum coherence' is defined by these analyses in two related ways:

(i) appropriate sequencing of material which follows the internal logic of a subject domain; and

[79] Ibid.

(ii) all materials, staff development, learning programmes and assessment emphasising the same aims and learning objectives in a mutually-reinforcing way.

While our evaluation has detected unevenness in the way in which these forms of coherence are displayed in individual subjects within the European Baccalaureate, the framework as a whole emphasises a form of curriculum coherence which is detectable in the commitments of staff and the reactions of students. As analysts criticise the performance of national systems in respect of curriculum coherence[80], this feature of the European Baccalaureate should be seen as a considerable asset.

Discussion around the different models presented in this section will need to take due account of such considerations of curriculum coherence. It is important to preserve the values identified by Interparents in their communiqué of February 2008. "The European Schools concept engenders citizens with an open-minded, multi-lingual and multi-cultural background, thus constituting an educational model which many schools might consider an enrichment of their national curricula." [81]

5.9 Introduction of new technology into test taking and test marking

5.9.1 Overview

Many have observed that new technology must be considered to offer solutions to current difficulties experienced by the European Schools in the operational and logistical processes involved in managing paper-based examinations. This section considers those suggestions and makes recommendations for future development.

This observation has been particularly in respect of:

- on-line test taking; and
- on-line marking

5.9.2 On-line Test Taking

The first proposal made by many Chairs is that a student's work that contributes to A and B marks could begin to be taken on-line. We agree with this approach and consider it unwise to introduce on-line assessment into final written examinations without having used earlier assessment sessions to pilot the effectiveness of the approach. There are a number of issues which need to be considered carefully before a move to encouraging students to take tests on-line.

- **Appropriateness for all examined subjects.** 'Essay' responses lend themselves very well to being completed on a personal computer. Those subjects which require calculations, technical drawing, the construction of graphs and diagrams are less well served, as of course are Art and Music.
- **Availability to all students**. Whereas it would be possible to decide that computers would be used by some subjects and not by others, it is not possible for some students within a subject to use a computer but not others if the overall assessment

[80] Ibid

[81] Interparents Memorandum, 27 February, 2008

goal of reliability is to be achieved. It is an important principle that all candidates should take an assessment under the same conditions (duration of examination, invigilation procedures etc.) and the introduction of computers for some is likely either to advantage or disadvantage a sub-section of the cohort.

- **Security**. The need to ensure that the computer contains no directories or drives which are likely to assist a candidate in examination challenges administration in respect of security.
- **Impact**. The introduction of computers for test-taking should be accompanied by classroom use. Teaching and assessment modes should be closely matched.
- **System cost**. The price of low-memory laptops and the prevalence of USB devices increasingly makes possible the use of technology.

5.9.3 Marking on-line

The second dimension of a 'technology' solution concerns the use of script marking on-line. A suggestion from a Chairman's report in 2004 proposes that scanned images should be distributed digitally to examiners to prepare before going to Brussels. A more radical suggestion is that it is possible for examination scripts to be scanned and for digital images to be transmitted to a central system from which examiners are able to download student work for marking. The proposal is, we believe, that schools would carry out their own scanning.

We suggest that pilots of this are carried out to identify the range of issues involved. It has been our own experience that there are many quality issues concerned in schools direct involvement in scanning. This stems from the need to 'track' individual script images (papers) throughout the process via the use of bar coding. The financial investment in scanning technology of the type required in schools would seem disproportionately expensive for equipment which was used infrequently.

The Cambridge approach has been to work with volume scanners to which student scripts are sent. Appropriate volume scanning bureaux exist across Europe.

The scope of e-marking can cover paper-based examinations, oral and aural tests or examinations taken at a computer.

There are typically considered to be a number of benefits of working within an e-marking system.

Quality benefits. The system allows for multiple marking of the same script, for the sending of previously marked scripts to check that a marker continues to apply the appropriate standard, for test exercises to be run to ensure that a marker has acquired the necessary competence and that all marking decisions are accurately recorded in real-time providing instant feedback on marking quality on item level data if required.

Efficiency benefits. The system removes many of the pressures of manual script handling. It enables access to a wider pool of markers and removes the stage of manual data entry which can sometimes introduce clerical inaccuracies into the process. The reliability benefits of marking online can provide the necessary tools to monitor and control the marking process.

A number of providers have developed online marking solutions. We know of Viatique in France, and in England both DRS and RM have developed systems to meet the needs of examination boards.

In this section we describe the Cambridge experience of working with one such provider, Research Machines plc.

5.9.3 Recommendation

We suggest that a pilot activity be designed first relating to students' work in class. A study exploring the feasibility of sending work between schools for cross-moderation of marking standards would provide information on students' usage, administrative time involved and the utility of moderating across schools.

The study, once evaluated, could then be extended to one subject in Part B examinations, before subsequent ramp-up to the final written examinations.

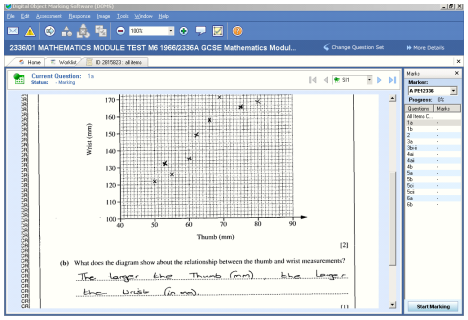
5.9.4 Marking On-Line - The Process Flow

Typical Script Packet

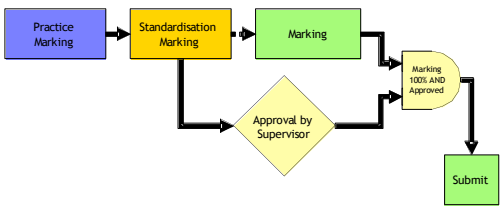
Processed by scanning bureau

Scripts are despatched from schools to a scanning bureau. In order to prepare the scripts for scanning it has been necessary for answer booklets to be printed with bar code labels and clear 'no write' areas to facilitate clean scanning. Scanning organisations have two complex things to achieve: management of the inventory to process scripts in the right order to achieve as much efficiency as possible and capturing of the data itself.

DOMS displays candidate 's work



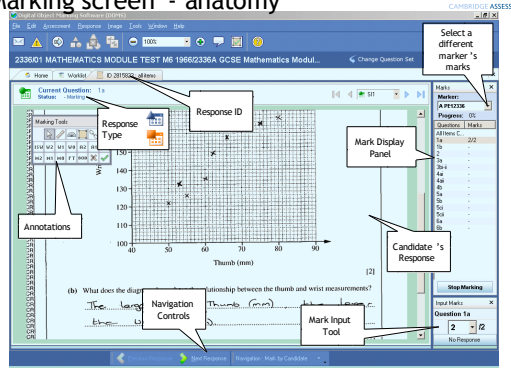
The DOMS Process Overview



Process-sequencing is less of an issue with the scale of the European Schools but in the Cambridge model it provides the opportunity to ensure that large volume examinations can be prioritised over those with fewer entries.

The bar coding identifies not only details regarding the origination of the script but also what is referred to as 'Product Reference Data' (PRD). The PRD will identify the structure of the question paper at item level and the maximum mark for each item. It then allows automatic calculation of marks.

Marking screen - anatomy



In the system employed by Cambridge Assessment, digitised images are presented on request to an examiner. This means that marking speed is increased. Candidate scripts are, of course, anonymously presented.

The annotations permitted by the system enable records of examiners' judgements but digitised images can be presented with or without annotations for training and review purposes.

An important part of the e-marking process is the design of a standardisation model in which definitively marked scripts are presented in both marked and unmarked mode enabling markers to test the accuracy of their marking on a 'trial' script.

Communications can be sent between markers, so that queries between raters can be clarified swiftly. It is also possible to present markers with a section of a question paper so that only certain questions are marked by a 'specialist' marker.

The costs of managing the system are difficult to quantify. Negotiations with suppliers will lead to variations in the split between operational fixed overhead costs and variable costings. The development in Cambridge was driven by quality rather than financial benefits. However, having worked with the system for 4-5 years in Cambridge Assessment as a whole, it is now possible to point to significant efficiencies in terms of speed of processing and quality improvements through improved reliability, and also financial savings.

The financial savings are being made in terms of reduced costs associated with the face-to-face meetings that it is now possible to avoid.[82] However, it is necessary to say that savings are achieved by virtue of the volume of script images that are processed through the system – currently just over 3 million scripts a year.

[82] Residential marking has not been our practice. However, meetings to ensure a common marking standard and to review marking have taken place on a face to face basis.

It is not obvious that savings can be achieved irrespective of the volume of scripts processed. Examiners have responded very well to the introduction of on-screen marking but it has so far been used widely for assessments at 16 for which response styles are likely to be more structured and less lengthy than for the 'essay' style of question more frequently used in assessment at 18.

5.9.4 Recommendation

We suggest that an investigative visit be made to an examination board that employs marking on-line strategies and that the system in use is appraised against requirements for the European Baccalaureate

5.9.5 Examiners' experience – marking on-line

Assessment literature suggests that the transition from paper to computer-based marking cannot be taken for granted from the marking point of view. Comparability between the two modes needs to be established. On-screen assessment can inhibit the reading comprehension of judges and this is particularly likely to be the case with the 'essay' responses of European Baccalaureate assessments written in 3 to 4 hour examinations.

A study carried out by Cambridge Assessment[83] has paid attention to examiners' cognitive load whilst marking in each mode. Measuring mental, physical and temporal demand, performance effort and frustration, data suggests that on-screen marking is more cognitively demanding than traditional paper-based marking due to examiners experiencing a heightened level of frustration during on-screen marking.

It may well be the case that a considerable amount of such frustration stems from the novelty of the experience, the inherent responsibility involved in marking and the characteristics of the software.

As software develops and examiners become more used to the system, cognitive frustration will reduce. Screen reading exacts a greater cognitive load on examiners and on-screen navigation is considered by examiners to be more challenging than paper navigation. A number of examiners have suggested that having an overview of the script, as is possible with paper, makes for greater confidence in the consistency of their marking. Examiners also reported that they were less likely to access other scripts when marking on screen. This factor might be considered important given the way that examiners use comparison within their judgement making processes.

The mean marking time per page on screen is greater than on paper, but this difference is not significant. Given the reduced administrative activities that examiners would be expected to complete, it is possible that marking on screen becomes a quicker activity than marking on paper.

Overall, the move to marking on screen must be regarded as a radical and high-risk development if the European Baccalaureate is to retain the services of existing examiners. It might be considered more appropriate to move in smaller steps:

- Conducting a cost-benefit analysis of the use of remote marking using paper scripts.

[83] On-screen Essay Marking, October 2008, Johnson and Nadas.

- Using scanning bureaux and on-line marking with one or two subjects to assess feasibility.
- Continuing to monitor a technology environment which changes rapidly. New 'players' in this environment are e-ink providers and pen technologies.
- Investigating other uses of technology which may be considered to yield more immediate cost savings without heavy upfront infrastructure investment.
- Alternative approaches to oral assessment, for example, might be considered in which all orals are recorded and moderated by external examinations.

5.9.5 Recommendation

Consideration should be given to the development of a project plan by which a move to on-line marking might be effected over a period of three to five years.

5.10 Costs in the European Baccalaureate

5.10.1 Overview

The cost analysis of the 2008 Report has provided an extensive view of financial spend in the European Baccalaureate.

Overall, operating costs of the European Baccalaureate have increased by 17.6% between the 2006 and 2008 session. There has been only a 3% increase in candidates and many fees have been held without inflationary increase over this period.

It is clear that a situation in which costs increase at a rate greater than the increase in candidature is neither sustainable nor an appropriate platform on which to base future expansion.

5.10.2 Establishing a benchmark

Cost comparisons between different organisations are difficult. The profile of actual and absorbed costs will vary and constrain direct comparison. The costs identified in the Annual Report on the European Baccalaureate relate largely to what might be termed 'direct costs' – they are actual costs expended in the delivery of examinations rather than overhead costs. To establish a 'like-for-like' benchmark, we have conducted an analysis of all examination-related costs across University of Cambridge International Examinations at a level equivalent to the European Baccalaureate.

On average, for each syllabus in our benchmark analysis, direct examination costs account for 36% of the examination fee to the candidate. A category referred to as 'other direct costs' covers those costs not easily assignable to a particular syllabus and represents 9% of the examination fee. Staff costs account for 20% and a 'contribution to overheads' in respect of premises, costs, IT infrastructure, human resources, finance and research and evaluation accounts for a further 25%. A target 10% of the examination fee is reserved for qualification and service development and contribution to public benefit spend in the University (overseas scholarships etc).

Table 5.2 shows this diagrammatically.

It should be remembered that the category of direct examination costs is an average; some syllabuses will incur costs in excess of the average.

CIE - Profile of Costs and Overheads as proportion of Examination Fee
Direct Exam Costs **36%**
Other direct costs **9%**
Contribution to Overheads: Staff costs **20%**
Contribution to overheads: Premises, Information Technology, Human Resources, Finance, Research and Evaluation **25%**
Contribution to Qualification and Service Development, to Educational Investment and to University Scholarships and benefits **10%**

Table 5.2 Benchmark Information

Exemplar Examination Fees – UK A Levels and IB Diploma
A typical single subject examination fee for UK A level ranges from £76 to £82 (on the basis of the awarding body Edexcel)
£94.00 most Language subjects UK A Level (Edexcel)
£97.20 Science subjects UK A Level (Edexcel)
A typical IB Diploma subject fee £50, plus a student registration fee £73.

Table 5.3 Benchmark Information

5.10.2 (a) Recommendation

We recommend that the Baccalaureate Unit establishes appropriate benchmarks with other examination boards by which it can appraise whether its costs represent value for money.

The 2008 Report to the Board of Governors calculates that on the operational costs expended on the Baccalaureate, an average per capita cost would equate for the 2007-08 session to a figure of € 655.49. Although this is not levied as a 'fee' it provides a base for discussing the level of operating costs incurred by the European Baccalaureate.

Although a little outside the fee charged by Edexcel or the IBO, it is not significantly so.

A candidate taking 4 A Levels and 2 AS subjects with Edexcel might incur a fee of £466; and for the IB Diploma £376. CIE fees are kept to a minimum because of its not-for-profit status, and relationship with the university. As such its fee structure does not bear comparison. To

both the IB and Edexcel fees must be added a Centre registration fee of £2500 - £5000 per annum. In this context, European Baccalaureate "costs per candidate" do not look out of line.

There is one very significant caveat, however. In the case of the examination boards used as a comparator with the European Baccalaureate (IBO/Edexcel) the fees are only in part made up of direct costs whereas for the European Baccalaureate such a "fee" if charged would represent 100% of costs. There would be no contribution to overheads, payment for staff and development. If a fee at this level is charged for students in Type III schools, it must be suspected that a degree of cross-subsidisation would occur from the overhead charges paid by Type I and Type 2 Schools through existing subsidies. Further expansion at this rate would become a 'cost' to the system.

CIE Analysis of Direct Cost Expenditure	
Setting Costs - 13% of all direct costs	Setting costs for 3 question papers (we do not operate with single component question papers for school-leaving examinations)
Question paper production - 18.25% of all direct costs	This includes the cost of question paper management – the cost of QP evaluation, of print layout, illustrations, proof reading and printing in-house under secure print conditions. Economies of scale achieved through large volume production.
Marking costs - 29.42% of all direct costs	Costs of marking papers
Standardisation costs - 12.82% of all direct costs	Costs of initial meetings and activities designed to ensure that marking standardisation has been completed.
Awarding costs - 11.4% of all direct costs	Costs of marking review, re-marking all candidates on grade thresholds and where performance on different components is misaligned
Script scanning - 4.5% of all direct costs	Applied as appropriate

Table 5.4 Benchmark Information

[Costs taken to provide benchmark information include those of greatest relevance to the European Baccalaureate. This, together with some rounding effects, means percentages do not total 100%]

Direct costs do not include:

- Despatch costs
- Warehousing

- Other staff costs indirectly involved with the process.

Table 5.4 provides a budget by budget analysis of expenditure under the heading of CIE direct costs. Although this model is an 'average' based on CIE's own processing volumes, it provides a tool by which the costs incurred by OSGES might be evaluated. It must be emphasised in this comparison that CIE costs account for 36% of an examination fee which is one of the lowest-cost charged.

The proportion produced for the direct costs spend of the European Baccalaureate is significantly different.

Table 5.5

European Baccalaureate Cost Profile	
Setting Costs	**23.59%**
Translation costs	**4.5%**
Print costs	**1.1%**
Marking costs (orals and Brussels)	**59.3%**
Awarding costs	**11.4%**

5.10 (b) Recommendation

We recommend that the Baccalaureate Unit produces a profile not only of direct costs but also of indirect costs and overheads in order to consider the business model which would fund further expansion of the European Baccalaureate.

5.10.3 Differences in cost profile – European Baccalaureate and benchmark

Notwithstanding the necessary caveats, it is apparent that there are certain budget headings where differences in the proportion of spend are of interest:

- A greater proportion of spend (23.69%) is spent on setting European Baccalaureate question papers than those of CIE's. This is surprising because it takes no account of the costs incurred in respect of the time spent by teachers as part of their contracted time in providing question paper proposals.
- In respect of European Baccalaureate setting costs, 12 subjects incur higher costs than the average of CIE setting costs at syllabus level. It should also be borne in mind that in the case of a CIE syllabus this represents three question papers and not one.
- When question paper production costs (i.e. all the quality assurance processes required to take the question paper to print stage) are added to setting costs, the CIE proportion of spend is 31.25% against the Baccalaureate spend of 28.1% (taking

setting and translation costs together). However, the CIE figure includes staff costs. We do not believe that the European Baccalaureate cost includes the costs of staff in the BAC unit.

- Marking costs account for 29.4% of CIE direct costs; 42% if standardisation costs are taken into account. However, they account for 59.3% of European Baccalaureate total costs. This is very significantly greater. The CIE percentage is of the direct cost figure rather than of total costs as with the European Baccalaureate.

- Awarding costs (i.e. the quality assurance costs of marking review and remarking candidates in particular categories) are 18.24%. The nearest equivalent cost for the European Baccalaureate might be considered to be the overall quality assurance function delivered by the Chairman at 11.4%.

The current level of marking cost in the European Baccalaureate has two major implications:

- examination fees must be kept at a level that proves prohibitive to the wider adoption of the European Baccalaureate

- the high level of marking cost constrains development in other areas: the review and development of syllabuses, investment in new technologies, and the like.

Marking costs are incurred by examiners visiting schools during oral examinations and by the residential marking exercise in Brussels.

5.10.3 Recommendation

We recommend that the Baccalaureate Unit reviews in particular the costs incurred in marking and question paper setting where costs are higher than for the benchmark.

5.10.4 Oral examining costs

The oral examinations are arranged in the candidate's school. Topics and questions are designed by the candidate's teacher. They are not always provided for the examiner prior to the examination. On the basis of the 2008 review of oral examinations, external examiners consider themselves less involved in the assessment process than the candidate's teacher. It is proposed that an alternative approach to oral examination be considered in which a teacher conducts an examination which is recorded either for sound alone or sound/video, and the recordings sent to the external examiner for appraisal and marking. Whereas typically an examination board would require a moderation sample, the importance of oral assessment within the European Baccalaureate would suggest that all examination recordings should be sent for a second external marking.

This proposal also offers a timetabling advantage in that the European Schools might wish to consider at what point oral assessment might most appropriately be held. Certainly the timing should allow for a process of cross-moderation across oral examinations from all schools. Such an approach to oral examination is likely to achieve considerable cost-savings. Given that it would also avoid the problem of overload in participating in a large number of orals in a short period, the approach may well be considered to offer some advantages in terms of quality.

5.10.4 Recommendation

We recommend that a review of the external examiner's actual attendance at all oral examinations be conducted. It incurs high cost. It also introduces significant constraint into the flexibility of the timetable.

5.10.5 Presentation of question papers

It may be possible to review the costs associated with the Chairman's sign-off of all question papers (€8832.33). There may well be many other functions of this meeting in March, which it would be unfortunate to lose, but in terms of question paper sign-off alone, we consider that this might be done by a review of submitted question papers with accompanying documentation and the opportunity for the Chairman to arrange subsequent communication with Inspectors where necessary. A Chairman's comment makes clear that not all papers are ready for sign-off at this date and a different approach to the process might assist this factor.

5.10.5 Recommendation

We recommend that a review of the March presentation of question papers to the Chairman by the Inspectors be conducted with a view to reducing cost by a different process.

5.10.6 Cost of language papers

The practice of compiling all papers for a single language in a common meeting is a sensible one and the costs of producing language papers is not significantly higher than for other subjects, except that the costs associated with the production of the German L1-L4 papers seem disproportionate. The German Committee held three meetings (September, December, January) whereas French held two and English two. On the information supplied it is not evident why this variation exists and there is a potential need to standardise the number of meetings and possibly the location.

5.10.6 Recommendation

Consideration might be given to standardising the number of meetings required in question paper consideration in order to control costs.

5.10.7 Cost of setting science papers

Each of Mathematics, Chemistry, Physics and Biology incurred setting costs greater than the CIE average. The attendance of all Inspectors/experts is the main driver for this expenditure. We appreciate that these subjects are taken in all schools and that a greater than average number of Inspectors might consider it necessary to be in attendance but we recommend review of this practice.

5.10.7 Recommendation

We recommend a review of the practice by which inspectors/subject experts from each country are required to attend question paper meetings.

5.10.8 Residential marking costs

The tables supplied in the section dealing with the financial aspects of the 2008 Report provide a very helpful analysis of the way in which, with the current marking approach, the cost of marking a script is a variable of the size of entry. The range spans from €12.05 per script for English L2 to €621 for Art or Spanish. All examination boards experience the phenomenon where some subjects are more expensive to mark than others (in CIE's experience it is most frequently Music where our costs might not be too out of line with the €170.46 cited in the report).

However, the costs associated with residential marking raise the question of value for money:

- when the script load is fewer than 10 scripts, the marking cost fee accounts for only 17.65% of total expenditure per examiner compared with 82.35% costs on travel and subsistence
- in only two cases (English L2 and Art in French) were the marking fees greater than costs of travel and subsistence.

Consideration might be given to a pilot in which a team of examiners working remotely (i.e. in their homes) and co-ordinated by a lead Examiner or Inspector began a non-live marking exercise in which they were asked first to mask trial scripts to establish a marking standard and then to mark a batch of paper scripts from a previous session. Marking outcomes and examiners' comments might be evaluated and the findings of this pilot be used to evaluate whether a change away from residential marking in Brussels would offer a quality advantage as well as a cost-saving.

5.10.8 Recommendation

We recommend a staged move away from residential marking. We propose the establishment of a pilot study to investigate the feasibility of conducting European Baccalaureate marking in examiners' homes.

6 Securing Quality in the European Baccalaureate

6.1 Overview

All examinations may be regarded as 'high stakes'. Each, whatever currency and recognition is applied by end-users, will have an impact on a student's motivation, learning and self-esteem. However, those examinations which mark the end of secondary education and provide for progression to university are of the highest importance to individuals and impose high standards of accountability on assessment bodies. Examination boards and assessment authorities need strong governance and control mechanisms in order to discharge their responsibility to users, and to safeguard standards of design and delivery at every stage of the assessment process.

In establishing such governance and control mechanisms, organisations turn frequently to quality systems that are used in a general management context, such as the ISO standards of performance management, or to specific Codes of Practice which bring together policies and procedures relevant to the delivery of high quality assessment.

In this section we review a range of quality assurance models and self-audit approaches likely to be of value to the European Baccalaureate in its future development and expansion:

- ISO certification and other quality assurance systems
- The establishment of an Assessment Code of Practice
- Models for monitoring and reporting compliance
- Internal self-audit
- Membership of associations related to educational assessment in Europe and internationally
- The development of Quality Assurance Training Programmes

6.2 ISO Certification and other quality assurance systems

6.2.1 Adoption of the quality standards enshrined in ISO 9001

While the establishment of a code of professional practice may be regarded as the cornerstone in defining appropriate standards of operation, it is a statement of those standards rather than a quality system designed to improve performance against the standard. A focus on improving performance against a defined standard of operation is a necessary element of achieving the ISO kitemark. The International Standards Organisation (ISO) in Geneva, Switzerland,[84] operates as a network of national standards institutions covering 148 countries. Its original focus was the development of technical and industrial standards. More recently, through ISO 9001:2000, it has moved into quality management systems applicable where an organisation:

- needs to demonstrate its ability consistently to provide products that meet user and applicable regulatory requirements; and

[84] www.iso.org

- wishes to enhance user satisfaction through the effective application of the system, including processes for continual improvement and the assurance of conformity with user and applicable regulatory requirements.

A number of examination boards, including CIE, have achieved ISO 9001 certification. Key quality management principles of user satisfaction, effective leadership and a systematic approach to management underpin the philosophy of the standard. Compliance with this standard provides confidence that the administration and management of an organisation meets best practice.

The process of obtaining certification is rigorous but the evaluative and self-audit approaches it requires leads to improvements in operational activities.

6.2.1 Recommendation

Consideration may be given to the adoption of ISO as a quality standard.

6.2.2 Other quality assurance models

The ISO standard is a generic one rather than one specifically designed for a certification organisation.

An earlier initiative to develop a quality standard designed specifically for organisations responsible for "the Certification of Personnel" led to the development of the European Standard, EN45013. Although more targeted than the subsequent standards enshrined in ISO 9001, the approach of EN 45013 provides little of practical implementation regarding the operational detail of certificating bodies.

Recognising the need for quality management systems which more helpfully meet the circumstances of examination boards, Cambridge Assessment has been involved in the development of its own quality assurance processes. The first example of this is the initiative taken in developing quality assurance procedures to be used by examining bodies in the area of language testing. In 1990 the Association of Language Testers in Europe (ALTE), was established to provide a forum for providers of language examinations for speakers of other languages. With 27 members across the European Union, and covering 23 languages, ALTE has played a role in the development of teaching and testing of foreign languages in Europe.

In 1994, members decided that it was essential to adopt a formal Code of Practice which "would both define the standards that current and future members would agree to aim to meet in producing their examinations and serve as a statement to consumers of those examinations of what they should expect".[85]

The broad areas of the ALTE Code of Practice have been redeveloped as detailed questionnaires in the form of checklists. Since 2001 these checklists have been used to monitor standards through self-assessment and informal peer monitoring.

The ALTE framework provides workable procedures and programmes of improvement. External monitoring is introduced as a secondary activity to confirm that minimum standards

[85] www.alte.org/quality_assurance

are being met rather than as a primary function at the beginning of the process to identify that they are not.

As a consortium of bodies operating in 23 languages, it is also essential that the ALTE framework should be flexible enough to recognise the diversity of institutions and contexts in which examinations are used.

The ALTE checklists are undertaken on a voluntary basis and cover a huge range of standards: as an example, an extract from the section on the "Test Construction" unit demonstrates this below:

Unit 1 - Test Construction								
Code of Practice	**Focus Points**			**Explanations and or references**	**Self-evaluation**			
		Yes	No		INP	UR	AD	GP
B. Test development, test construct & context. Routine procedures. B1. Explain how content and skills to be tested are selected (for each component of the examination).	i. Were you required to make reference to an official source in establishing the content of the examination, such as national curriculum guidelines, e.g. from a Ministry?							
	ii. Do you make use of domain or profile descriptions (skills and components)?							
	iii. Is use made of the Common European Framework of Reference? If yes, how?							
	iv. Were the specific assessment goals developed by your examination body itself? If so, how?							

6.2.3 Recommendation

Consideration may be given to a review of the ALTE standards and to the use of such an approach in the European Baccalaureate.

6.3 The establishment of a Code of Practice relevant to the organisation and mission of the European Baccalaureate

It is clear that the annual document *Arrangement for Implementing the European Baccalaureate*, produced by the OSGES, operates as an essential handbook within the European Schools. However, while it provides a valuable source of information it tends toward the definition of administrative procedures rather than professional standards.

The development of a document which codifies not only the organisation's policies and procedures but also its model of 'best practice' provides a valuable platform for discussion about quality. In England and Wales a Government regulator, the QCA, works with each of

the Awarding Bodies[86] to agree a Code of Practice harmonising best practice, which is reflected with only minor variations in Scotland, in international examining and elsewhere. Assessments within the United States are defined within a framework of standards issued by the American Psychological Association. Such documents provide a necessary framework of professional practice. A key objective of the UK Code of Practice is to ensure consistency between examination boards and across different qualifications. The Code is designed "to provide a basis for high standards in all aspects of the examination process and therefore to make an important contribution to upholding public confidence in these qualifications."[87]

The Code of Practice covers all aspects of assessment, from a statement of general principle through to (2005) detailed administrative arrangements. The structure of the UK Code of Practice is described in the table below:

Introduction	
Principles underpinning the code of practice	
Section 1:	Responsibilities of awarding bodies and awarding body personnel
Section 2:	Awarding bodies' relationship with centres
Section 3:	Preparation of external assessments including provisional mark schemes
Section 4:	Standardisation of marking: external assessment
Section 5:	Awarding, marking and moderation: internal assessment
Section 6:	Awarding, marking review, maintaining an archive and issuing results
Section 7:	Arrangements for candidates with particular requirements
Section 8:	Malpractice
Section 9:	Enquiries about results and appeals
Section 10:	Access to marked examination scripts

6.3 Recommendation

The Board of Governors of the European Baccalaureate Examination Board may wish to give consideration to a codification of current, and future, practice with the production of a European Baccalaureate Code of Practice.

6.4 Appointment of professional officers with subject responsibility

The range of activities covered by the Code of Practice highlights the breadth of activities involved in managing the curriculum, assessment and teaching dimensions of examination practice. Typically, a cadre of professional and administrative staff are employed by an

[86] The Awarding Bodies are Edexcel, AQA and OCR in England, WJEC in Wales.

[87] QCA (2003) Code of Practice.

examination board to ensure that the annual cycle of activities are managed in full compliance with the tenets of the Code of Practice. Such staff have either professional responsibility for a designated subject or operational responsibility for a specific function.

In the case of CIE, for example, such staff comprise respectively 34% Assessment Services and 34% Operations).[88] The Board of Governors of the European Baccalaureate Examining Board might wish to give some consideration to the possibility of appointing professional officers able to oversee such a range of functions. Given the scale of the organisation and the size of its current operation, a proposal to appoint three officers responsible respectively for the Sciences, for Languages, and the Humanities might be considered appropriate. This level of staffing would enable a number of quality functions which are not currently as well developed as would be considered advisable.

- Management of a regular process of syllabus revision and liaison with member States about necessary matriculation requirements, and with inspectors about the availability of markers
- Liaison with the Bac Unit in respect of question paper production
- Co-ordination of internal marking during Years 4 and 5, Year 6 and the first semester of Year 7
- In liaison with a subject's Lead Examiner the co-ordination of standardisation exercises for all external examiners prior to marking.

A particular focus of the work of the professional officer might be concerned with processes of cross-moderation between schools. Although this function is, to some extent, carried out by the Inspector with lead responsibility in a subject, the pressures of a busy caseload can squeeze the amount of attention that can be given to this role on some aspects of quality assurance between schools. In terms of the control of standards, ensuring that teachers from across the European Schools network provide a 'moderation' of the marks given in internal assessment is vital.

It would be advisable for such professional officers to have had teaching experience, and desirable for them to have had examining experience in their designated subject.

6.4 Recommendation

The Board of Governors of the European Baccalaureate Examination Board may wish to give consideration to the appointment of three subject professional officers to take responsibility for a range of assessment-related tasks in the Sciences, Languages, and the Humanities.

6.5 Models for monitoring and reporting compliance

The establishment of a Code of Practice provides the opportunity to identify the process by which performance will be monitored and compliance measured. It is apparent that evaluation processes are currently in use in connection with the European Baccalaureate. The report of the Joint Inspection approach conducted at Karlsruhe, the Evaluation of the 2008 Oral examinations and the Annual Report to the Board of Governors each, in their way,

[88] Other staff are employed in Divisions responsible for Syllabus Development (12%) and Development Activities (20%).

demonstrate the way in which evaluation is considered important. Developing a monitoring system, anchored to the goals of the Code of Practice, formalises such approaches into an evaluation loop by which future practice may be improved.

There is an opportunity to exploit the strong relationships that exist between schools in the European Schools network by the introduction of a number of monitoring approaches:

- double marking (on a sampling basis) at least the part B assessment of the preliminary examinations.
- Sampling internal assessment work (a moderation sample would represent a very small proportion of all scripts).

The focus on 'moderation' as a form of 'self-audit' recognises the value placed on internal assessment within the European Schools and strengthens the concept of shared values across the network.

6.5 Recommendation

The Board of Governors of the European Baccalaureate Examination Board may wish to give consideration to the establishment of a process of cross-moderation for A & B marks – both by teachers in other schools and by the Lead Examiner.

6.6 Internal self-audit

Self-assessment is a critical step towards achieving best practice. It brings many benefits to an organisation, including encouraging the participation and development of staff, and ensuring their commitment to organisational goals. Self-assessment against the QCA Code of Practice by all UK awarding bodies is, in fact, built into a separate QCA regulation which states explicitly that:

"An awarding body should carry out self-assessment to demonstrate that it is regularly and effectively evaluating its quality assurance systems and quality control arrangements for accredited qualifications against the regulatory requirements."[89]

Particular forms for such self-audit might be concerned with reviewing standards within schools' marking. In recognition of the value of an assessment–related self-audit system, CIE developed a programme, Quality Share, to assist organisations in identifying the areas of their work that might merit the focus of a self-assessment approach. The key areas of assessment practice it covers are set out in the following modules:

- Syllabus production
- Question paper/mark scheme production
- Marking
- Internal assessment and moderation
- Grading

89 Regulatory Arrangements for the Qualifications and Credit framework, August, 2008, Section 2.7

- Post-results services
- Arrangements for Special Consideration
- Security of examinations

6.6.1 The key features of the Quality Share self-assessment model

The self-assessment model provides an opportunity for staff at all levels to engage in constructive debate about key areas of progress in core examination activities. It can also be used to measure progress over a period of time.

The concern is not only for achieving the highest professional standards and practice, but also being able to *demonstrate* such achievement. Based on our established Code of Practice, we have followed a process that has enabled us to:

- identify and define the 'core' examining processes which are relevant to all bodies engaged in examination work
- break down the key constituent parts of such processes
- convert these parts into a series of critical questions which can act as standards of best practice
- present these questions in such a way that respondents can measure their current position in achieving these standards
- prepare users not only to indicate current progress, but identify supporting evidence which will demonstrate compliance or non-compliance.

An example of the approach can be provided by the section on moderation in which a series of detailed statements are provided against which actual performance can be evaluated. Throughout, an evidence-based approach is encouraged in identifying which descriptor is more appropriate.

Annex: Extract from CIE Quality Share section 4

Assessment criteria	Evidence/Comments	To be developed	Some Good Practice	Good Practice	Best Practice
		1	2	3	4
Internal assessment is carried out by teachers at a centre when they evaluate candidates' work against assessment criteria. The assessment criteria might take the form of performance descriptors or a mark scheme, and will be included in the published syllabus. Care must be taken to ensure that the assessment criteria are clearly worded and as unambiguous as possible, and that they demand a standard of work consistent with that required in the externally assessed components of the syllabus. To what extent are policies and procedures in place to ensure that...					
(a) when a syllabus is being designed the skill areas designated for internal assessment are most appropriately assessed in that way, and not by written question papers?					
(b) the weighting allocated to the internal assessment as a proportion of the overall assessment scheme for a syllabus is appropriate, and conforms to the aims and assessment objectives of the syllabus?					
(c) the construction and wording of the internal assessment scheme and assessment criteria are clear and unambiguous?					
(d) the range of activities covered by the coursework component and the number of skill areas and criteria within the scheme are sufficient to allow appropriate differentiation, without being burdensome for centres'					

There are many other benefits of a systematic and highly developed self-assessment approach, but perhaps the one factor often overlooked in the context of quality assurance processes generally is the fact that they generate considerable internal debate, and engage staff at all levels in focussing on performance measures and the possibilities for improvements to be identified and acted upon; the process itself has intrinsic benefits for the organisation and undoubtedly leads to greater motivation of staff, and improved channels of communication.

6.6 Recommendation

The Board of Governors of the European Baccalaureate Examination Board might give consideration to the adoption of a process of self-audit to continue to monitor performance and set standards for improvement.

6.7 Membership of associations involved in educational assessment

6.7.1 An International Association of Educational Assessment (IAEA)has been in existence for over thirty years. In 2000, a European Chapter was established – the Association of Education Assessment – Europe[90].

[90] http://www.aea-europe.net

The role of the AEA-Europe is to "help co-ordinate efforts to improve assessment systems and practice throughout Europe, through contact between organisations involved in examinations and assessment research and individuals working professionally in this area. These contacts promote scientific knowledge and the application of that knowledge to practical problems in school systems".[91] The AEA-Europe:

- is engaged in a programme which provides: a professional accreditation system, recognising the achievements of individuals and also providing a framework for professional development.
- provides opportunities for professional development through courses, workshops and other events.
- has a website at www.aea-europe.net which operates as a communication tool between members.
- facilitates members' participation in projects funded by the European Union.
- enables collaboration with other associations and agencies for which assessment represents at least a part of their interest.

6.7.2 The Cambridge Assessment Network

Cambridge Assessment provides an assessment network designed to create links between assessment professionals. It provides a programme of professional development in assessment and related issues through a comprehensive programme of conferences, seminars and courses. These activities are described on the Cambridge Assessment website at www.cambridgeassessment.org.

6.7 Recommendation

OSGES may wish to give consideration to membership of European and International Associations of Educational Assessment to create a forum for assessment discussion and the exchange of best practice.

6.8 The development of quality assurance training

Opportunities for training and development in assessment can be either:

- Open programmes delivered by examination bodies
- Bespoke programmes tailored to the specific requirements of the European Baccalaureate.

6.8.1 Open programmes delivered by examination bodies

CITO (the national institute for educational measurement in the Netherlands)

[91] http://www.aea-europe.net

- CITO offers courses that range from practical training support to educational workshops on the development of qualitative test methods and testing using computers.

Cambridge Assessment International Study Programme

- Examination development
- Public Examination Administration Certificate in the Principal and practice of Assessment
- Programme of regular seminars on assessment–related topics.

6.8.2 Bespoke programmes

A number of recommendations in this Report have pointed to specific opportunities for training:

- Training teachers who might be involved in writing assessment tasks
- Training teachers as assessors
- Training examiners

A wider programme concerned with tasks relevant to the operation of an examination board as a whole may also be considered. Such a programme might be used as a context to general production of the European Baccalaureate Code of Practice.

Recommendation 6.8

In parallel with the possible development of bespoke training and of examiners for teachers as item writers and assessors, we recommend a programme of training for all those involved with policy and practice in the European Baccalaureate.


PART OF THE
CAMBRIDGE ASSESSMENT
GROUP