DEFINING THE INFRASTRUCTURE FOR A NATIONAL ITEM BANK SERVICE

Niall Sclater, Eddie Boyle, Joanna Bull, Clive Church, Patrick Craven, Rowin Cross, Myles Danson, Leah Halliday, Ian Howie, John X Kelly, Steve Lay, Moira Massey, Mhairi McAlpine, Diane McDonald, Mary MacDonald, Steve Rogers, Su White

Defining the Infrastructure for a National Item Bank Service

Niall Sclater, Rowin Cross, Diane McDonald, John X Kelly: University of Strathclyde, Glasgow, Scotland
niall.sclater@strath.ac.uk

Eddie Boyle, Moira Massey, Leah Halliday, Steve Rogers: EDINA National Data Service

Joanna Bull: Eduology

Clive Church: Edexcel

Patrick Craven: OCR

Myles Danson: Loughborough University

Ian Howie: Stow College

Steve Lay: UCLES

Mhairi McAlpine, SQA

Mary MacDonald: COLEG

Su White: University of Southampton

Abstract

Item banks (sometimes known as question banks) have been around for many years but are not yet widely used in the UK. There are clear benefits such as economies of scale when items are built across a subject area or sector. When this is coordinated centrally items are more likely to be peer reviewed, validated properly and to adhere to technical, interoperability and accessibility standards. Quality can be enhanced by delivering the items to larger numbers of candidates, leading to improvements following analysis of item usage data. However there is currently no satisfactory way for these to be stored and made available to potential users; the available commercial learning object repositories are unable to deal with assessment content adequately.

In an attempt to solve such issues and to begin to define the infrastructure of a distributed national item bank service, the Item Bank Infrastructure Study (IBIS) brought together individuals and institutions in the UK with key expertise in areas relating to item banks. The study was funded by JISC

under the Exchange for Learning (X4L) Programme with financial contributions from three of the exam boards involved – Edexcel, the Scottish Qualifications Authority (SQA) and the University of Cambridge Local Examination Syndicate (UCLES). This paper extracts the key points and conclusions from the full report which can be downloaded from www.toia.ac.uk/ibis. An accompanying paper in these proceedings, *Conceptualising Item Banks*, defines items, item pools and item banks, and outlines the main components of a possible distributed item bank service.

1. Review of existing item banks

Eight organisations who owned item banks were surveyed during the study. Most of these were focussed on a specific subject, mostly in numerate disciplines though one contains items in most areas of the Scottish FE curriculum. All banks offered multiple choice questions. Some offered other objective item types such as multiple response and fill-in-the-blank. Only one included essay items – expected to be delivered on paper and marked manually. Intended usage of the item banks was split fairly evenly between formative and summative assessment. It was suggested that items intended for diagnostic assessment of A-level knowledge would also be suitable for formative and summative assessment in FE.

Staffing varied dramatically between projects. One item bank was looked after by only two people. Another project used fifty authors organised into thematic teams to develop items. The largest had 66 authors supported by a technical advisory group, validators, a steering group and an overall project manager. Training or guidelines were offered by most projects to item developers.

The number of items held in the banks varied between 500 and potentially 10^{25} runtime realisations based on the use of random variables in mathematical items. An item bank obviously tends to grow over time but funding is more critical than age in determining its size. One bank which was only a year old already had 3,000 items. They are stored in a variety of formats: all digital but mostly proprietary and only some using the internationally recognised IMS Question and Test Interoperability (QTI) format.

There is considerable variation in the internal structure and organisation of these item banks. Only two of them used aggregations such as assessments though half of them offered automated item selection. Most of them store some metadata to describe the items. Just two however used recognised international metadata standards or used elements beyond the immediate details of the item such as author name, copyright details etc. There would seem to be a clear need for the development of an effective metadata schema for classifying items.

No banks surveyed stored candidate data within the item bank. It will be essential to collect this data in the future in order to ascertain the quality of items and thus enhance them or retire them from the bank.

Items are gathered in a range of ways – from Microsoft Word to uploading over the Internet to submitting them on paper. All of the banks had some kind of quality assurance procedure in place though these varied considerably. Those used for high stakes assessment had stricter quality assurance process in place. Around half of the banks had some facility for item analysis but this was relatively underutilised as a means of enhancing items.

There were frustrations with the current software available; one project identified problems in transferring data between Word, statistical packages, databases and spreadsheets. Despite considerable amounts of time, effort and money, items are often trapped in a proprietary format, and their application beyond their immediate context is thus very limited. Furthermore the lack of adequate metadata to accompany the items means that they may never be discovered with a larger item bank.

2. Interagency Communication

Within the UK most of the initiatives concerning online assessment have been led from the individual educational sectors with a major contribution from the agencies and commercial organisations that work with them. Despite the benefits that can accrue from the shared use of interoperable items there has been little work performed so far to achieve this vision.

Within HE there is significant use of online testing. The availability of VLEs and simple CAA tools has enabled more practitioners to experiment with the use of online assessment to support their teaching. However without a common curriculum sharing of items has been limited. There has been minimal seepage of items to other sectors including Sixth forms and FE. Higher education has though been supported by a network of subject-based learning and teaching support services, the twenty-four LTSN (Learning and Teaching Support Network) Centres, now incorporated in the Higher Education Academy. Some of these have produced item banks for their subject areas.

The support network for FE colleges adopts a different approach, being regionally rather than subject focussed. Through funding from JISC there is now a network of Regional Support Centres (RSCs), which provide advice and guidance on the use of ICTs.

Apart from in Scotland (with the COLEG Online Assessment project – COLA) and with the recent Basic/Key Skills testing pilots there has not been a national drive to increase the uptake of online testing in FE. Within individual colleges and training centres (without sharable item banks) there are not the group sizes to provide the economies of scale that college based online testing can provide (unlike HE). Paradoxically, with a shared curriculum there are enormous potential benefits from sharing items through IBIS. There is

currently no mechanism (outside Scotland) to avoid nearly 500 colleges duplicating the development of items for the same qualification.

Becta is the English government's key partner in the strategic development and delivery of its ICT and e-learning strategy for schools and the learning and skills sectors. Becta has a pivotal position in supporting the drive to the adoption of interoperability and by implication facilitating the development of item banks and associated tools and systems.

The Learning and Skills Development Agency (LSDA) is a strategic national resource for the development of policy and practice in post-16 education and training. Dependent on research funding, it supports a range of initiatives (not just ICT-based) across the lifelong learning sectors. Staff are equipped to perform research and support around the implementation of item banks if required (and funded) and have expertise in interoperability standards.

The picture demonstrated so far represents a plethora of agencies (not all mentioned in this paper) funded from different sources within the four nations of the UK with disparate and often overlapping responsibilities that demonstrate little cohesion as far as item banks and associated tools are concerned. In fact, although there are still risks of parallel and unconnected developments, there are major recent and proposed developments that will facilitate improved communications.

4. Legal Issues

Certain legal preliminaries need to be attended to in considering the developing of a national item bank service. Ownership and provenance are perhaps the most obvious. Ownership rights issues arise both with the content coming in and also with the content going out at the other end.

Factors which necessitate our examination of the law and how it applies to the activities of a national service include:

- The protection of its reputation
- Regulation – activities attract legal compliance, for example data protection compliance
- Enforcement of rights – intellectual property
- Self policing – the ebay illustration
- Consumer Protection

For the service to be successful the content and the delivery mechanisms must engender trust. Thus reputation could be said to be the responsibility of all those who are engaged in any way with the service. The law can be used to protect that reputation in a number of ways including for the users by ensuring that what is promised is what is delivered and for the depositors that their rights are protected.

Whether or not any financial transactions take place with users the service should plan to treat customers legally. Existing consumer protection law, including that on the sale of goods and misleading advertising, applies online. Although users are unlikely to be considered consumers within the traditional definition it is suggested that in the interests of good customer relations practices and procedures which mirror consumer protection legal requirements should be adopted.

Consortium agreements are important for a number of reasons including for example the outlining of responsibilities of each party to the agreement. Details such as the technological specifications demanded from participating repository institutions can be spelled out here. It may be that minimum requirements (eg 99% up time for connection) are included in such contracts.

Institutions and individual authors will already have existing assessment content which they may have been using for years which when adapted would be very suitable for inclusion in an item bank. This historical content may in some ways be the most problematic in terms of ownership.

The service will generate usage data and user access data. The extent of this information can vary but technically it is possible to trace almost every transaction carried out by any party who browses or purchases content. This should clearly be anticipated in the infrastructure build and it is possible to establish ownership of this in favour of the item bank owner at that stage.

Because the service is likely to require individuals to register their personal information as a preliminary step to their use of the item bank certain legal obligations are imposed by the Data Protection Act 1998. Institutions are also bound by the disability legislation and many are working to comply with the requirements to make education materials accessible to all. As part of the infrastructure build, the service should create policies and procedures which are in line with the standards acceptable to disabled users.

The Freedom of Information Act 2000 gives a general right of public access to all types of recorded information held by institutions and as it is likely that such as service will be a recipient of public funding it will attract the scrutiny which the Act brings.

A commitment at start-up to protect the rights of the individual and an ongoing awareness of the possible pitfalls may be sufficient at this stage for the service to comply with the Human Rights Act.

5. Item analysis

Statistics have an important role to play in the development of a high quality item bank. They can for example be used to ensure the quality of questions and help remove bias due to factors such as gender and ethnicity. Students with difficulties can identified more easily and appropriate remedial assistance offered.

There are two major statistical theories in assessment: classical test analysis and latent trait theory. Classical analysis takes the test as its frame of reference rather than the item. It produces statistics which are relevant only to that test and to the group of students which attempts it. Latent trait theory on the other hand aims to produce statistics which are universal and applicable to all who attempt an item. There are two main forms of this theory: Item Response Theory and Rasch Modelling. Both claim to produce statistics which are independent of the students who take them and the items which are taken by them.

Item banking has in the past been associated with items which unfairly discriminate based on sex, ethnicity and first language. The test should of course be fair and discriminate only on the basis of the candidate's ability. Bias statistics are therefore often calculated to help identify items where candidate performance varies by group.

The following parameters may be helpful in identifying items of appropriate quality to be included in IBIS. In general, extremes of difficulty should be avoided. It is undesirable to include items which are not likely to provide much information about the candidate's ability. In order to maximise the information available about candidates, the discrimination of items should be as high as possible.

Careful use of the statistics generated can also help inform educational management decisions by providing concrete data on how candidates are performing. It can influence the adaption, reform and development of courses, by highlighting areas of study which are complementary as well as curricular areas which do not sit so well with the rest in a particular program. Statistical analysis can also help direct resources where they can be utilised most fully.

Assessment statistics can be monitored as students progress through a course, providing early indicators of any areas that students are finding difficult, and identifying students that require additional support. There are a variety of ways that these students can be identified, but usually those who have the lowest mean score (classical) or lowest ability (latent trait) will suffice. A more developed version of this would be to divide the course into the core and the extensions and concentrate on those who are finding difficulty with the core and thus might require additional help.

Grouping questions within criteria can identify candidates who are not performing adequately on one part of the curriculum – a notional 'pass-mark' can be set for each area of the course which can be used as a benchmark to identify candidates who are struggling with that area. These candidates could then be offered tailored feedback or additional support.

For the purposes of generating statistics, secondary metadata or usage data (structured information about the use of an educational resource) is essential. The IMS QTI v2.0 specification contains guidance on the recording of usage data. Usage statistics are dynamic, context-specific information, and are thus

held separately from static metadata. An item's metadata is not used solely for the discovery of the item, but also to facilitate the exchange of items between repositories. Usage data may well be proprietary and confidential, and would need to be dissociated from the rest of the item's metadata if the item is to be exchanged outside its current repository.

The structure envisaged by IBIS includes both banks held by IBIS and the infrastructure to exchange items between item banks and delivery systems held by separate organisations. For the IBIS banks, item analysis will take place on the items contained within the banks, and provide the associated usage data. For items which are held outside IBIS, the usage data will be sent to the item bank owner and a receipt sent to the brokerage system.

6. Metadata and searching

Many of the issues surrounding metadata for assessment are those facing cataloguers of learning resources in general. Best practice in learning object metadata creation, maintenance, quality assurance and searching has been developed by a large and active community of practitioners, and should be adhered to wherever possible. Searching technologies and architectural frameworks for assessment metadata should similarly adhere to common practice in order to maximise interoperability and availability.

However, currently available metadata standards, specifications and application profiles are generally inadequate for the cataloguing and discovery of assessment resources. The two major metadata standards, the Institute for Electrical and Electronics Engineers (IEEE) Learning Object Metadata (LOM) standard, and the Dublin Core Metadata Initiative standard contain only minimal reference to assessment objects, and certainly do not describe them with the degree of detail necessary to maximise the discovery and reusability of assessment resources. Application profiles of these standards, such as the IMS Question and Test Interoperability (QTI) version 2.0 LOM application profile and the TOIA/COLA assessment metadata schema, have attempted to address these problems, and offer first steps towards the development of a standardised element set and vocabularies for assessment objects.

Many of the issues surrounding metadata for assessment are those facing cataloguers of learning resources in general. Best practice in learning object metadata creation, maintenance, quality assurance and searching has been developed by a large and active community of practitioners, and should be adhered to wherever possible. Searching technologies and architectural frameworks for assessment metadata should similarly adhere to common practice in order to maximise interoperability and availability.

The UK LOM Core represents the results of real-world application of the IEEE LOM standard within the UK educational context. Developed through consultation with a broad community of practitioners, it identifies and codifies common practice to define a minimum common core of LOM elements and the values or content such as vocabularies associated with each element

which should be implemented. It defines mandatory, recommended and optional elements, and contains explanatory notes to the profile which reflect implementation experience with the LOM standard. The TOIA/COLA Assessment Metadata Application Profile is an application profile of the UK LOM Core itself, and was the first LOM application profile which attempted to define metadata for assessment in any detail. It was produced as part of the COLA project, an initiative to develop item and assessment banks for further education.

The recent revision of the IMS Question and Test Interoperability (QTI) specification resulted in a new QTI metadata category, *qtiMetadata*, containing ten new elements to describe the structure and behaviour of an item, and the nature of the interactivity it enables.

Other than the TOIA/COLA assessment metadata application profile, there is very little innovation in assessment metadata evident in current practice. In the IBIS survey only half of respondents appeared to understand what was meant by the term 'metadata', a result which is surprising and which is unlikely to be consistent with other online educational repositories. The low level of awareness of metadata, and in particular of metadata standards, among respondents is surprising, and illustrates the need for professional cataloguers to work alongside resource creators in order to maximise the reusability of assessment materials. Within small-scale item banks, browsing may suffice for discovering usable resources but within IBIS's Brokerage System, which aims to harvest metadata in order to provide users with a range of choices for download and purchase, resources which lack adequate metadata will be hidden and inaccessible.

7. Service delivery

Detailed service delivery requirements will depend on the service scenario(s) adopted, the types of assessments being delivered by the service, the usage models adopted and whether or not sales of items via the service will be offered. As an example, the resilience required of the service will be variable according to whether the service offers on demand summative assessments for awarding bodies, such as en masse A Level scheduled assessments. Further to this, aspects of service delivery will be dependent on the technical requirements of the hardware/software configuration to be implemented and this detailed information could only be presented after a technical specification for the system had been completed.

The technical, organisational and legal aspects of IBIS will differ depending on what type of assessments are offered. Four different scenarios are envisaged: formative, diagnostic, local summative, external summative and vocational/workplace assessments.

The primary purpose of formative assessments is to promote learning by providing feedback; it does not count towards the final mark of a module. The main beneficiary under this scenario is the student. Feedback is clearly an

important feature in questions used for formative purposes. Such assessments tend to be lower risk scenarios though service provision must of course be robust. It is desirable that such assessments be available as close to 24/7 as possible. Evidence suggests that given the choice students tend to access such material outside of standard office hours. These are tests which aim to determine a student's prior knowledge of a subject area. The benefits of this type of assessment are both to the student and the instructor. Results of such assessments are often more than simply a final score. Students are informed of their strengths and weaknesses and often directed to appropriate learning material in order that they may improve. The general issues mentioned under the formative assessment section apply.

Summative assessments count towards the final mark of a course. The stakes are therefore high but of a sliding scale from 100% of overall credit downwards. Thought should be given to determining and policing a set of minimum standards that must be met prior to local delivery of national item bank content for summative purposes. In the case of multiple/national summative assessments procedures must be in place to ensure the integrity of the test content is not compromised at any local assessment facility. Any live data feed to and from a remote local assessment centre and a national question bank needs to be robust, secure, and strictly controlled. Summative assessments generate the highest load on a delivery system as all students start at the same time, and many finish at the same time.

External Summative Assessments differ from Local Summative Assessments only in that they are delivered away from the chief stakeholder.

Vocational/workplace assessments differ in that the delivery area may be outside an educational institution/facility or a recognised testing centre. Recent developments have seen assessments being delivered directly to the workplace through a web browser. Less commonly for perhaps more specialised subjects and skills the use of Personal Digital Assistants (PDAs), or personal wireless enabled laptops may be supported.

The IBIS service provider should be responsible for:

- The day to day running of hardware and communications equipment
- Raising fault reports with the relevant supplier for any faults and escalating problems as appropriate
- Taking regular file store backups and placing backup tapes in a secure place
- Maintaining the operating system and database software
- Database and file store administration
- Development of any interfaces as appropriate in line with user requirements, unless this work is done by the provider of the systems

A helpdesk should be provided to act as the primary point of contact for all enquiries concerning IBIS. Support teams at local institutional level are likely

to be required, which can provide assistance to users within their own institutions. Local support teams might include, for example, site representatives, expert users, learning technologists, ILT champions and/or IT support people. Existing channels of support within institutions should be utilised. This model has worked well for other complex JISC services.

8. Security, access and authentication

The success of any CAA exercise depends on its effectiveness in measuring a candidate's responses against the learning objectives of the programme. While the accuracy of the assessment and its appropriateness to the learning objectives are of concern, it must also be ensured that the assessment is fair; it should be impossible to copy or have existing knowledge of either questions or answers, be it through tampering or inadvertent exposure. Without such assurances, the assessment will not necessarily be an accurate measure of a student's abilities. Protocols of use and security of CAA software are therefore important.

At the heart of the additional security issues is the need to minimise misuse or inadvertent exposure of items across a distributed and unsynchronised infrastructure. 'Traditional' security issues which occur in assessment and the underlying computer systems must of course also be addressed.

Security is essentially concerned with limiting the risk of malicious and inadvertent interference with service, a lapse of which may cause loss of revenue and customers or exposure to legal redress.

Arranging for a more able student to take a test using a candidate's identity is known as misrepresentation. In diagnostic and formative testing misrepresentation may not be considered a major issue as the candidate is only 'cheating' him or herself. Care must be taken however if completion of the assessment is part of a learning requirement.

Another variant on misrepresentation is the potential for a remote person to view data being rendered to the screen and send back the answers to the candidate's computer using 'control taking' software. The only way to avoid this is to ensure that the access controls policy on the candidates' computers is closely managed and all potential security holes plugged. Detailed logs should be kept for access to all assessment computers, the computers 'cleaned' after each high-stakes test and regular security audits carried out before each test.

Items may be exposed during authoring, validation and management. Good working practices must be put in place to ensure that items and solutions are not exposed during these processes; paper copies must be kept under lock and key, authoring and validation should be carried out in private and once an item is accepted, all unencrypted copies, both paper and electronic destroyed. Items may also become exposed during purchasing, and test compilation, marking, when not automatic (e.g. essay type answers) and quality control. Purchasers (customers), markers and quality controllers must follow

procedures similar to above. In the case of high-stakes pools, it may be necessary to restrict potential purchasers to browse sample questions only.

The main source of item exposure is of course during testing itself. If the test is delivered to all candidates at the same time and individual items will not be reused at a later date then delivery exposure does not matter. If however reuse of items or whole tests may occur then use of items must be carefully controlled otherwise copying and statistical analysis may enable prediction of a large enough percentage of the test format, effectively defeating the test.

CAA in general tends to reduce the incidence of copying, as candidates can be provided with different items from a test pool, delivered either randomly or depending on previous answers, or the tests may be delivered at different locations or times. Even if the same item is delivered to all candidates in a test, then the questions are typically only available for a short period of time and the candidate's response is visible on the screen for an even more limited period, making copying of responses difficult.

The distributed nature of CAA as envisaged in the national service means that there are many potential points of disruption. Most vital, because of its time dependency and high operational costs will be actual assessment delivery. This susceptibility to service denial is one of the main drivers for constructing a service so that items are not directly served from an item bank during a live assessment. Tests should thus be 'pre-compiled' and served directly from a logically separate test delivery system.

While the central purpose of a national service is to unlock institutional item banks for sharing, the inherent distributed nature of its infrastructure results in many additional implications for both item exposure algorithms and security in general over traditional CAA. On the whole however it is anticipated that these can be acceptably minimised for the majority of assessments undertaken.

Conclusion

This paper has:

- Examined some of the item banks already in existence
- Discussed some of the players with interests in item banks
- Presented the main legal issues involved in setting up a national item bank service
- Shown some of the benefits of using statistics to analyse items and improve item banks
- Described how metadata can be used to facilitate item description and retrieval
- Looked at issues around service delivery, security, access and authentication

In short, the paper has discussed many of the issues it would be necessary to consider in the development of a national distributed item bank service. The

full IBIS report expands on all of these and other issues. There is now the required expertise in the many areas surrounding item banks and service delivery in the UK to make the development and provision of an efficient and effective national service viable.

References

Barker, P., Cross, P., Fernandez, A., Hardy, S., Jewell, N., Poulter, M., Powell, A. (2003) *RDN/LTSN resource type vocabulary*, version 1.0, Resource Discovery Network. www.rdn.ac.uk/publications/rdn-ltsn/types

Bennett R. (2001) *How the Internet will Help Large-Scale Assessment Reinvent Itself*. Education Policy Analysis Archives Vol.9 No.5

BSI - British Standards Institution (2002) *A code of practice for the use of information technology in the delivery of assessments*

BSI - British Standards Institution (2002) *A code of practice for the use of information technology in the delivery of assessments*

Bull J. and McKenna C. (2000) Computer Assisted Assessment Centre (TLTP3) Update *4th International Computer Assisted Assessment Conference* Loughborough 21st-22nd June

Bull, J. and McKenna, C. (2003) *Blueprint for Computer-assisted assessment*, Routledge Falmer.

Campbell, L.M. (2004) *UK Learning Object Metadata Core* draft 0.2, University of Strathclyde, Glasgow www.cetis.ac.uk/profiles/uklomcore/uklomcore_v0p2_may04.doc

Carlyle, A. (2001) 'Developing organized information displays for voluminous works: a study of user clustering behaviour', *Information Processing and Management*, 35, 5 (July), 677-699

Case, S.M., Swanson, D.B. (2001) *Constructing Written Test Questions for the Basic and Clinical Sciences* third edition, National Board of Medical Examiners, Philadephia

CETIS (2001) CETIS *Reference: metadata* Bolton: CETIS. www.cetis.ac.uk/encyclopedia/entries/20010905154527

Clarke, S., Lindsay, K.,McKenna, C. and New, S. (2004), 'INQUIRE: A case study in evaluating the potential of online MCQ tests in a discursive subject', ALT-J 12.3, forthcoming.

Cliff, Pete (2002) *RDN resource types*, version 1.0, UK: Resource Discovery Network. www.rdn.ac.uk/publications/cat-guide/types

Cocks, Grant (2002) Crosswalks between e-GMS, LOM and the Curriculum Online metadata scheme UK:Simulacra Media.

Conole, G. (2004) *Assessment as a catalyst for innovation*

Crocker L. and Algina J. (1986). *Introduction to classical and modern test theory* NewYork: Holt, Rinehart and Winston.

Currier, Sarah and Jane Barton (2003) 'Quality assurance for digital learning object repositories: how should metadata be created?', in Cook, J and D McConnell (eds) *Communities of Practice: research proceedings of the 19th Association for Learning Technology Conference* (ALT-C 2003), Sheffield: Association for Learning Technology, pp 130-142.

Currier, Sarah, Jane Barton, Rónán O'Beirne and Ben Ryan (2004) 'Quality assurance for digital learning object repositiories: issues for the metadata creation process', *ALT-J: Research in Learning Technology* 12 (1). metadata.cetis.ac.uk/guides/files/currbartobeiryan_altj_6.doc

Darimont, A. W. (2002) *Environmental Scan of Pricing Models for Online Content, Report II, Business Models for Object Repositories.* www.edusource.ca/craw/RPT_Pricing_Model201.htm

Data Protection Act (1998)

Davey, T. and Nering, M. (2002) 'Controlling Item Exposure and Maintaining Item Security' in *Computer-based Testing: Building the Foundation for Future Assessments* (Eds: Mills, C.N., Potenza, M.T., Fremer, J.J. and Ward, W.C.)

Davey, T., Miller, T.R. (1992) 'Effects of item bias and item disclosure on adaptive testing.', Annual meeting of the Psychometric Society, Luneburg, Germany

Davies, P. and Lewis, S. (2004), 'Computerised peer-assessment that supports the rewarding of evaluative skills in essay writing (CAP) and programming (Coursemarker)', *Proceeedings of the 8th International Conference on Computer Assisted Assessment*, 85-90.

DTI (2000) 'The Consumer Protection (Distance Selling) Regulations 2000 (Summary)', the DTI Consumer & Competition Policy, www.dti.gov.uk/ccp/topics1/guide/distsell.htm Accessed 24/08/04

Dublin Core Metadata Initiative www.dublincore.org

Fan, X. (1998) *Item Response Theory and Classical Test Theory: An Empirical Comparison of their Item/Person Statistics* Educational and Psychological Measurement Vol. 58 No. 3

Frary R.B. (1988) *Formula scoring of Multiple choice Tests (correction for guessing): Instructional Topics in Educational Measurement* Educational Meaurement: Issues and Practices Vol. 7 No. 2

Friesen, Norm, Jon Mason and Nigel Ward (2002) 'Building educational metadata application profiles', *Proceedings of the International Conference on Dublin Core and Metadata for e-Communities*, 63-69.

Friesen, Norm, Sue Fisher and Anthony Roberts (2004) *CanCore Guidelines for the Implementation of Learning Object Metadata (IEEE 1484.12.1-2002)* version 2.0 Athabasca: Athabasca University. www.cancore.ca/documents.html

Gibson W. and Weiner J. (1998) *Licensing and Certification Test Construction: A balancing act.* CLEAR Exam Review, Vol 7 No 1, Winter

Goldstein H. and Woodhouse G (1996) The Statistical Analysis of Institution Based Data, in Goldstein H and Lewis T (eds) Assessment: Problems, Developments and Statistical Issues. J Wiley and Sons

Gorissen, Pierre (2003) *Quickscan QTI: usability study of QTI for De Digitale Universiteit* Utrecht: De Digitale Universiteit. www.gorissen.info/Pierre/qti/Quickscan_QTI_UK.pdf

Greenberg, Jane (2003) 'Metadata generation: processes, people and tools', *Bulletin of the American Society for Information Science and Technology,* January, 16-19.

Gulliksan H. (1950) *Theory of mental tests* New York, Wiley

Hartman, Lee (2003) 'No sound change in LCSH', online posting to HISTLING mailing list, 23 September. listserv.linguistlist.org/cgi-bin/wa?A2=ind0309D&L=histling&P=R2

Herd, G., Clark, G. (2003) *Question Types*, Glenrothes College.

IEEE LTSN: IEEE Learning Technology Standards Committee (2002) *Draft Standard for Learning Object Metadata*, IEEE 1484.12.1-2002, New Jersey: IEEE Standards Department. ltsc.ieee.org/wg12/files/LOM_1484_12_1_v1_Final_Draft.pdf

IMS (2001*) Using IMS Content Packaging to Package Instances of LIP and Other IMS Specifications* version 1.0 implementation handbook, US: IMS. www.imsglobal.org/implementationhandbook/imspack_handv1p0.html

IMS (2003a) *IMS Learning Information Package Accessibility for LIP* version 1 final specification, US: IMS. www.imsglobal.org/accessibility/index.cfm

IMS (2003b) *IMS Digital Repositories Interoperability – core functions best practice guide* version 1.0 final specification, US: IMS. www.imsglobal.org/digitalrepositories/index.cfm

IMS (2003c) *IMS Content Packaging Information Model version* 1.1.3 final specification, US: IMS. www.imsglobal.org/content/packaging/index.cfm

IMS (2004) *IMS Question and Test Interoperability: Item Metadata and Usage Data*, version 2.0 public draft, US: IMS Global Learning Consortium. www.imsglobal.org/question/index.cfm

IMS (2004b) *IMS Question and Test Interoperability: Item Information Model* v2.0 public draft, US: IMS. www.imsglobal.org/question/index.cfm

IMS (2004c) *IMS Meta-data Best Practice Guide for IEEE 1484.12.1-2002 Standard for Learning Object Metadata*, version 1.3 public draft, US: IMS. www.imsglobal.org/metadata/index.cfm

IMS (2004d) *IMS AccessForAll Meta-data Specification*, version 1.0 base document, US: IMS. www.imsglobal.org/accessibility/index.cfm

JANET-CERT (2004a) 'Protecting Computers', www.ja.net/CERT/JANET-CERT/prevention/machines.html. Accessed 10/06/04

JANET-CERT (2004b) 'Protecting Networks', www.ja.net/CERT/JANET-CERT/prevention/networks.html. Accessed 10/06/04

JISC (2001a) *Copyright and Licensing Guidelines* www.jisc.ac.uk/index.cfm?name=projman_copyright

JISC (2001b) *Senior Management Briefing Paper 13: Developing an Information Security Policy* www.jisc.ac.uk/index.cfm?name=pub_smbp_infosec

JORUM (2002) *Draft JORUM Depositor Licence Agreement* www.jisclegal.ac.uk/publications/johncasey_apends.htm

JORUM+: The JORUM+ Project Teams at EDINA and MIMAS (2004) *JORUM Scoping and Technical Appraisal Study Volume 5: Metadata*, UK: JORUM. www.jorum.ac.uk/research/publications.html

Letson, Russell (2001) ‘Taxonomies put content in context’, Transform (December). www.transformmag.com

Maurer, Donna and Todd Warfel (2004) ‘Card sorting: a definitive guide’, *Boxes and Arrows*, April. www.boxesandarrows.com/archives/card_sorting_a_definitive_guide.php

McAlpine M. (2001) *BluePaper 2: A summary of methods of item analysis.* CAA Centre, University of Luton

McCabe M. and Barrett D. (2003) *It’s a MUGS Game! Does the Mathematics of CAA Matter in the Computer Age?* Maths CAA Series, April 2003, LTSN Maths, Statistics and Operational Research Network, University of Glasgow.

Miller, Paul (2001) ‘Architects of the information age’, *Ariadne*, 29. http://...

Morrison (1972) *A Method for Analysis of Choice Type Question Papers* in Nuttall D. L. and Wilmott A. S. *British Examinations – Techniques of Analysis*, NFER, Windsor

Mozley, H.N., Whitely, G.C. (1977) *Law Dictionary* 9th edition, Butterworths
Powell, Andy, Mikael Nilsson, Ambjörn Naeve and Pete Johnston (2004) *DCMI Abstract Model* US: Dublin Core.

Ricci, Christian (2004) 'Developing and creatively leveraging hierarchical metadata and taxonomy', *Boxes and Arrows* (May). www.boxesandarrows.com/archives/developing_and_creatively_leveraging_hierarchical_metadata_and_taxonomy.php

Roszkowski, Michael and Christopher Lukas (1998) 'A distributed architecture for resource discovery using metadata', *D-Lib Magazine* (June)

Rowley G. L. and Traub R. E. (1977) *Formula Scoring, number right score and test taking strategy* Journal of Educational Measurement Vol. 14 No. 1

Sclater, N. (2003) *TOIA-COLA Assessment Metadata Application Profile* version 1.2, University of Strathclyde http://www.cetis.ac.uk/profiles/uklomcore/toia_cola_metadata_v1p2.doc

Sclater, N., Howie, K. (2003) User requirements of the 'ultimate' online assessment engine, *Computers and Education*, Vol 40, No 3, Elsevier Science Ltd, 285-386

Sclater, N., MacDonald, M. (2004) Putting interoperability to the test: building a large reusable assessment item bank, ALT-J, *Research in Learning Technology*, Vol 12, No 3, September 2000

Sclater, Niall and Phil Barker (2004 – forthcoming) 'Using Learning Object Metadata in online assessment'

SHERPA (2004) *The Move Towards Open Access of Research Output: Briefing Paper* www.sherpa.ac.uk/documents/open_access_briefing2.pdf

SHERPA (2004) Working advice on a deposit licence for E-Prints (SHERPA internal discussion document)

Shmoo Group, 'How to Write Secure Code', www.shmoo.com/securecode/ Accessed 24/08/04

St Pierre, Margaret and William P LaPlant (1998) *Issues in Crosswalking Content Metadata Standards*, National Information Standards Organisation. www.niso.org/press/whitepapers/crsswalk.html

Stage, C (2003) *Classical Test Theory or Item Response Theory, The Swedish Experience*, EM 42, Centro de Estudios Publicos ISSN 1100-696-X

Tennis, Joseph T (2003) ‘Data collection for controlled vocabulary interoperability – Dublin Core audience element’, *Bulletin of the American Society for Information Science and Technology*, December/January, 20-23

Thelwall, M. (2000) 'Computer-based assessment: a versatile educational tool', *Computers & Education* 34 pp37-49; - haven’t quoted from this but should!

UK Government BusinessLink E-commerce guide www.businesslink.gov.uk/bdotg/action/layer?r.l1=1073861197&r.s=tl&topicId=1073866263 Accessed 22/08/04

Zakrzewski, S. and Bull, J. (1998) 'The mass implementation and evaluation of computer-based assessments', *Assessments and Evaluations in Higher Education*, 23(2) 141-152