

## ENGINEERING A TESTBANK OF ENGINEERING QUESTIONS

*Su White<sup>1</sup>, Kate Dickens<sup>2</sup>, H C Davis<sup>3</sup>*

**Abstract** *e3an is a collaborative UK project developing a network of expertise in assessment issues within electrical and electronic engineering. A major focus is the development of a testbank of peer-reviewed questions for diagnostic, formative and summative assessment. The resulting testbank will contain thousands of well-constructed and tested questions and answers. Question types include objective, numeric, short answer and examination. Initially a set of metadata descriptors were specified classifying information such as subject, level and type of cognitive skills being assessed. Academic consultants agreed key curriculum areas (themes), identified important learning outcomes and produced sets of exemplar questions and answers. Questions were authored in a word processor then converted to the database format. Major pedagogical, organisational and technical issues were encountered, including specification of appropriate interoperable data formats, the standards for data entry and the design of an intuitive interface to enable effective use.*

**Index Terms** *Assessment, IMS, Interoperability, Metadata, QTI, Testbanks, Usability*

### BACKGROUND

The Electrical and Electronic Engineering Assessment Network (e3an) was established in May 2002 as a three year initiative under Phase 3 of the Fund for the Development of Teaching and Learning (project no. 53/99) The project has been led by the University of Southampton in partnership with three other UK south coast higher education institutions; Bournemouth University, The University of Portsmouth and Southampton Institute [1]. The project is collating sets of peer-reviewed questions in electrical and electronic engineering which have been authored by academics from UK Higher Education. The questions are stored in a database and available for export in a variety of formats chosen to enable widespread use across a sector which does not have a single platform or engine with which it can present questions.

We were aware of the potential difficulties associated with recruiting adequate numbers of question authoring consultants to produce well-formed questions and populate a demonstration version of the database. We deliberately chose to use a word processor as our authoring tool since we assumed it would present the smallest barrier to question production. We made a detailed specification of the contents

of the question, and provided authors with word processor templates which we planned to automatically process and convert into the data items in the question database. If necessary authors could provide clear hand written detailed feedback in the form of tutor's notes to be scanned.

Authors were left to their own devices to write the questions, and we assumed that the peer review process would identify and remedy any gaps in the quality of the question. We also assumed that this laissez faire approach would enable us to better specify and explain requirements to future authors. However, when we came to transferring the electronic versions of the completed into the database, the number of issues we encountered were significantly larger than we had originally envisaged, and to an extent which would not be sustainable in the longer term.

### PROCESS OVERVIEW

Activities associated with the production of the question database were divided into a number of distinct stages – the pre-operational stage, the collection of questions themselves, and then the post processing of the questions. We introduce the phases here and then describe them more fully in the subsequent sections of this paper.

The pre-operational stage involved the project team in answering some design questions:

- *Database contents and function and specification.* What formats did we wish to store in the database and how would we wish to access and use that information?
- *Target themes.* What themes or subjects within the EEE domain did we wish to collect questions about?
- *Metadata.* What descriptors about the questions did we wish to collect?
- *Question Templates.* How would we get EEE academics, some of whom may not have regular access to latest computing technology, to enter their questions? Once these questions had been resolved, we set about:
- Building theme teams by recruiting and briefing academic consultants
- Authoring and peer reviewing the initial set of questions

These activities were all carried out in parallel with the continuing specification and design of the question database. When the authoring task was complete, additional processing needed to be carried out by the core team to:

- process sample questions for inclusion in the e3an testbank

<sup>1</sup> Su White, University of Southampton, Intelligence Agents and Multimedia, ECS Zepler Building, Southampton, SO17 1BJ saw@ecs.soton.ac.uk

<sup>2</sup> Kate Dickens, University of Southampton, Intelligence Agents and Multimedia, ECS Zepler Building, Southampton, SO17 1BJ kpd@ecs.soton.ac.uk

<sup>3</sup> Hugh Davis, University of Southampton, Intelligence Agents and Multimedia, ECS Zepler Building, Southampton, SO17 1BJ hcd@ecs.soton.ac.uk  
e3an is funded under the joint Higher Education Funding Council for England (HEFCE) and Department for Employment and Learning in Northern Ireland (DEL) initiative the Fund for Development of Teaching and Learning (FDTL), reference53/99

## Session 5, Stream Number 4

- partially automate the conversion process
- review of the success of the authoring process
- identify good practice for future authoring cycles

### DATABASE CONTENTS AND FUNCTIONS

The testbank was conceived as a resource to enable academics to quickly build tests and example sheets for the purpose of providing practice and feedback for students, and also to act as a resource for staff providing exemplars when creating new assessments, or redesigning teaching.

Because questions for the database were to be produced by a large number of academic authors, we chose a simple text format as the data entry standard. In addition, in order to enable the widest possible use across the sector, it was important that questions from the bank could be exported in a range of formats

The question database was designed to allow users to create a 'set' of questions which might then be exported to enable them to create either paper-based or electronic formative, summative or diagnostic tests. Output formats specified included RTF, for printing on paper (as the lowest common denominator), html for use with web authoring, and an interoperable format for use with standard computer based test engines which is a subset of the IMS Question Test and Interoperability (QTI) specification [2].

There are many types of objective question; QuestionMark [3], one of the leading commercial test engines supports 19 different types, and the latest IMS QTI specification (v 1.2) lists 21 different types. However, the project team felt that they would rather experiment with a small number of objective question types, and decided to confine their initial collection to multiple choice, multiple response and numeric answer.

In addition to objective questions, the project team wished to collect questions that required written answers. We felt that some areas of engineering, design in particular, are difficult to test using objective questions. Furthermore, a collection of exam style questions, along with worked answers, was felt to be a useful resource to teachers and students alike.

### INITIAL SPECIFICATIONS

The four initial themes chosen were:

- Analogue Electronics;
- Circuit Theory;
- Digital Electronics and Microprocessors;
- Signal Processing

These subjects were seen as being core to virtually every EEE degree programme, and were chosen to reflect the breadth of the curriculum along with the teaching interests of the project team members. A theme leader was appointed for each subject area, drawn from each of the partner institutions. A particular additional concern was that the material produced should, as far as practicable, be

relevant to courses in electrical engineering. An electrical engineering subject specialist was therefore appointed to work with the four theme teams and encourage consultants to reflect "heavy current" interests.

### METADATA

Metadata is data that describes a particular data item or resource. Metadata about educational resources tends to consist of two parts.

The first part is concerned with classifying the resource in much the same way as a traditional library, and has such information as AUTHOR, KEYWORDS, SUBJECT etc. and such metadata has for some time been well defined by the Dublin Core initiative [4].

The second part is the educational metadata, which describes the pedagogy and educational purpose of the resource, and standards in this area are only now starting to emerge as organisations such as IMS, IEEE, Ariadne and ADL have come to agreement. For example, the user might see the ARIADNE Educational Metadata Recommendation [5] which is based on IEEE's Learning Object Metadata (LOM) [6].

The project team was keen not to allow the project to become bogged down with arguments about which standard to adopt, or to overload their academics with requirements to provide enormous amounts of metadata, so instead decided to define their own minimal metadata, secure in the knowledge that it would later be possible to translate this to a standard if required. The following list emerged.

TABLE I  
INITIAL METADATA SPECIFICATION

Descriptor	Contents
Type of Items	Exam; Numeric; Multiple Choice; Multiple Response
Time	Expected to take in minutes
Level	Introductory; Intermediate; Advanced
Discrimination	Threshold Students; Good Students; Excellent Students
Cognitive	Knowledge; Understanding; Application; Analysis;
Level	Synthesis; Evaluation
Style	Formative; Summative; Formative or Summative; Diagnostic
Theme	Which subject theme was this question designed for?
Sub themes	What part of that theme?
Related	What other themes might find this question useful?
Themes	
Description	Free text for use by people browsing the database
Keywords	Free text; for use by people searching the database. e.g. mention theorem tested

A set of word processor templates were created which allowed the question to be specified and to incorporate the e3an metadata [7]

We felt that the DISCRIMINATION field improved on the DIFFICULTY field found in the standard metadata sets in that it could be related specifically to the level descriptors used in the UK Quality Assurance Agency's (QAA) Engineering Benchmarking Statement [8]

### BUILDING THEME TEAMS

Consultants were recruited from the partner institutions. Project team members were initially asked to identify and canvass potential contributors from their own institution for each of the four subject areas.

Prospective consultants were then invited to attend a half-day training session. Individual briefings were organised for prospective consultants who were unable to attend on either one of the two dates offered.

The briefing session was divided into two main sections; an introduction to the e3an project and objective testing, and a meeting between the theme leader and members of the theme team to discuss and agree specific objectives for their subject theme area.

The introductory component included:

- A description of the e3an project, objectives, participants, timescales and deliverables;
- Overview of issues in student assessment, including the benefits of timely formative feedback and the outcomes-based approach to assessment advocated by the Quality Assurance Agency for Higher Education [9,10];
- Familiarisation with guidelines for writing effective objective test questions, question types: multiple choice, multiple response, numeric answer and text response.

The latter, practical components of the briefing, were considered essential since the design of effective objective test questions is an acquired skill [11]. Some general guidelines were presented [12], along with additional examples drawn from the electrical and electronic engineering curriculum. Specific examples demonstrated how indicative questions from a “traditional” examination paper might be converted into objective test format, and how an author might design a question to test a specific learning outcome.

The second part of the briefing session involved members of theme teams meeting with the theme leader to discuss and agree the key curriculum areas within their themes and the spread of question types that would most usefully support these sub-themes. This activity was conducted over the two separate events, with details finalised by email. Theme leaders initially proposed the main sub-themes or topics and their indicative level. Questions were classified as “Introductory”, “Intermediate” and “Advanced”; levels which may broadly correspond to the three levels of a full-time undergraduate programme in electrical/electronic engineering. However the schema is flexible enough to cope with the specialist nature of some degree programmes and acknowledge that both timing and intensity of study may vary between institutions.

There was also some debate amongst members of the project team about the designation of materials as being relevant the fourth year of an MEng programme. An additional metadata item of tutor information provides an

opportunity for question writers to append explanatory notes if they feel this will be appropriate.

### AUTHORING AND PEER REVIEW

Once the training and allocation of questions was completed authors took a few weeks to produce their questions and then met again to peer review each others questions, before the questions were accepted for publication in the database. Consultants were paid on successful peer review of their allocation of questions.

### DATABASE ORGANISATION

Once the questions were delivered, there were two questions which the project team needed to resolve; what information to store in the database and what database architecture to use.

We had already agreed that users might require the output in the following formats:

- Rich Text Format – in order to allow the creation of paper-based tests, and that there would be two presentations, namely the questions alone and the questions with worked answers.
- HTML – to enable questions to be put on the WWW, again in question only and question with answer format.
- XML to the Question & Test Interoperability (QTI) specification which will allow the questions to be used by a range of proprietary test engines including Blackboard and version 3 of Question Mark Perception.

We also acknowledged that many users might wish to import the questions into earlier versions of Question Mark, which do not support QTI, and for this reason we decided that we should also provide QML output, which is Question Mark’s proprietary XML format.

We wrote a programme to read the Document Object Model (DOM) of the questions that had been provided in the Word templates, and from this to create the XML QTI version of the question.

In theory it would have been adequate to store only the original Word documents as all the other formats could have been created dynamically from these documents. However speed limitations persuaded us that it was best to batch process the input documents and to store all the possible output formats.

The question of what architecture to use was decided by the fact that we needed to be able to distribute the database to individual academics, maybe working without system support; we wanted to be able to send them a CD and them to be able to install and query the database on their own.

We decided that this would be easiest using Microsoft Access, and that we would deliver the system packaged with the Access runtime for those that did not have Access on their machines. In parallel we have developed an SQL server version of the database, so that it can be queried over the

Internet. We may also consider making this version available to institutions for Intranet use.

### Question Retrieval and Export

The user may assemble a question set either by browsing the whole database and making individual selections, searching using individual or multiple keywords or by a combination of both these methods. At this stage the user may view the descriptive metadata for each question and may also preview the tutor version of each question in Microsoft Word.

Given the ubiquity of online shopping it was decided to use a trolley metaphor to enable the building of question collections. If the user wishes to go on to make use of a question or set of questions this is achieved by adding the item(s) to the "question trolley". The "question trolley" may be viewed at any time during the process and questions may be added/deleted. The question set may then be sent to the "question checkout" to be exported. The user is given an option of saving the question set so that it may be modified or used again at a later date.

At the "question checkout" the user may choose which combination of versions and formats the questions are to be exported in. Each question has a unique identifier and the suffix of each question identifies its version and format. Currently (summer 2002) each question can be found as an individual file after being exported individually. Future versions will also enable optional export of concatenated collections of questions in the various formats.

### The User Interface

Users are expected to follow a basic four-step process.

- searching/browsing the database,
- selecting a set of questions
- saving the set of questions, or adding the selection to a previously saved set
- exporting the questions in one of the available formats.

In order for the user to keep track of their progress through the system a "breadcrumb trail" is displayed which shows each stage in order and highlights the process currently being undertaken.

We first produced a fully featured working prototype which was modified after an intensive evaluation cycle (detailed further below). Initially the main menu provided the following options:

- opt to either create a new collection of questions;
- browse through the question bank;
- view/amend previous question collections;
- exit the question bank.

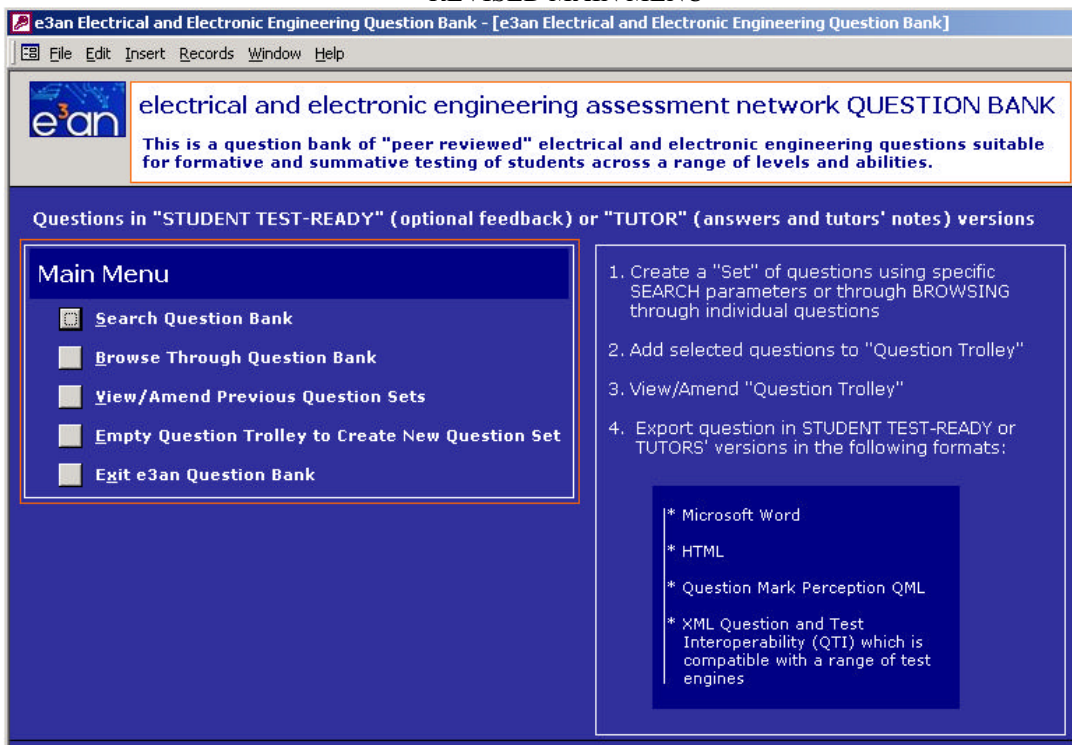
The example below (figure I) incorporates some additional features added after the initial usability trials

The search interface allowed the user to search by:

- subject theme;
- question type;
- level (i.e. introductory, intermediate, advanced).

The search results were displayed indicating the search parameters and links to both the question with the question description and keywords and the question metadata.

FIGURE I  
REVISED MAIN MENU



There were also options to select questions and then add them to the question trolley and also to search for additional questions. The question trolley interface was similar to the search results interface but it allowed the user to remove questions if necessary and it also provided an option to "save the question collection" and to "proceed to the question collection checkout". The "question collection checkout" interface was designed to allow the lecturer to first select the export style; tutor, student or both and then the export format (RTF, HTML, QML, or XML(QTI)).

### USABILITY TESTING

Three different methods of usability testing were used all of which may be classified as "discount usability engineering" approaches [13]:

- Scenarios – where the test user follows a previously planned path
- Simplified thinking aloud – where between three to five users are used to identify the majority of usability problems by "thinking aloud" as they work through the system
- Heuristic evaluation – the test user (who in this case requires some experience with HCI principles) compares the interface to a small set of heuristics such as ten basic usability principles.

The majority of the test users were lecturers who were asked to follow scenarios and/or the simplified thinking aloud method. Those who had experience of HCI methods (whether a lecturer or not) were initially asked to complete a heuristic evaluation in order to take advantage of their expertise. **Evaluation**

The project is taking an integrative approach to evaluation [14] and plans to review and revise the current implementation based on actual use. However in the earliest stages the focus was to evaluate usability. The first prototype was tested by three academics from UMIST and Manchester Metropolitan University a further six evaluations were carried out at Loughborough University using the three methods of discount usability engineering detailed above. Feedback was collated and the most obvious usability problems were fixed before the second prototype was tested at the partner institutions by a further twelve subjects.

The feedback was then collated again. A questionnaire following Nielsen's Severity Ratings procedure [15] was sent to three project members (two academics and one learning technologist) indicating where the problems existed, their nature, and any actions that had so far been taken to overcome them. These individuals were asked to rate the usability problems for severity against a 0 – 4 rating scale where 0 = I don't agree that this is a usability problem at all and 4 = usability catastrophe: imperative to fix this before product can be released; the mean of the set of ratings from the three evaluators was used to identify the most critical

usability problems which were then fixed and the third prototype was produced.

Prototype three was given an in-depth heuristic evaluation with an HCI expert. Again problems were identified and the prototype modified once more. The fourth prototype was tested using scenarios and the simplified thinking aloud methods within the team which resulted in the fifth prototype which is the version which has been released for general usage.

### Issues Identified

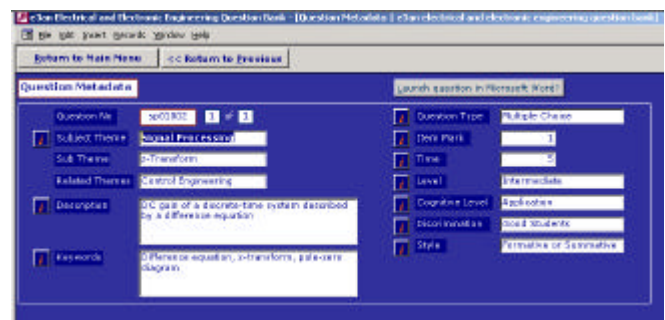
The evaluation process identified a number of issues when users accessed the database.

- Users were unsure what was meant generally by the term metadata;
- Users did not always understand the purpose of specific metadata items (e.g. question level);
- Inconsistent naming conventions within the retrieval process created confusion;
- Users found different stages of the process indistinct;
- Users were sometimes presented with more information than they needed.

### Explaining Meta Data

When browsing or after searching the testbank, users were given an option to view question details before selecting and saving. A key aspect of the database is the inclusion of metadata associated with each question. The usability trials showed that although users understand the concept of providing structured information describing the form and contents of the question database, the majority of academics are not yet familiar with the term metadata. Consequently an additional information icon was added next to metadata buttons which pops-up explanation as the mouse passes over the icon. The button display was changed from 'metadata' to 'View question metadata?' as shown in figure III. A similar issue was identified associated with the metadata contents. Information buttons were added to the specific metadata items to provide explanations of terms used. See figure II below

FIGURE II  
INFORMATION BUTTONS ADDED TO METADATA SCREENS





### What do you call question collections?

There was originally some ambiguity in the naming conventions used to describe question collections. The term proved confusing, perhaps because it was alliterative (e's and q's sounding the same) long-winded, and it appeared difficult to recall. The term was simplified to question set.

### Keeping track of progress

Although a breadcrumb trail had been included to assist users in keeping track of their progress, there were inconsistencies in the naming of processes on screen and on the breadcrumb trail, nor was it clear where along the trail a user was currently situated. The trail was changed to ensure that it matched terms used elsewhere for search and retrieval. A highlight was added to indicate to users their current progress through the trail. The screen displayed after a search and when a user was viewing the question trolley were very similar. Screens were numbered to reinforce the feedback on progress and an orange question trolley icon was added to the screen as shown in figure III.

FIGURE III

TROLLEY SCREEN SHOWING BREADCRUMB TRAIL, PROGRESS NUMBER AND TROLLEY ICON



Users found the process of exporting questions and saving a question set confusing. A dialogue box was added to prompt the user to save the question set before they proceeded to export.

### CONCLUSIONS

Subsequent testing via the scenario and simplified thinking aloud suggests that the usability problems originally found have been solved without creating additional problems. Most problems identified in each prototype were found by the first five test users, testifying to the power of these methods in identifying usability problems easily overlooked by users involved in the database specification. It shows how an "easy to use" database finding unexpected user problems. Unfortunately there is not yet any similar common methodology for the collaborative creation of database contents. For e3an the objective of achieving a high level of voluntary involvement among question authors was equally important to creating a useful testbank [16]. The time needed to manually process the first batch of questions far exceeded

our original estimates, and authors variously showed imagination in the variety of problems they created for us. However we have now incorporated our learning from that experience into our briefing sessions, a detailed guidance document, and a validation tool. We continue to use a word processor and template, and are actively working to share our learning with others in the HE community.

### ACKNOWLEDGEMENT

We would like to acknowledge the hard work of our colleagues in Southampton and our partner institutions; our initial group of authors, and all those who contributed to the evaluation of early versions of the database.

### REFERENCES

All url's last accessed July 2002

- [1] e3an Additional project details at <http://www.ecs.soton.ac.uk/e3an>
- [2] Smythe C, Shepherd E, Brewer L and Lay S "IMS Question & Test Interoperability: ASI Information Model", *Final Specification, Version 1.2, IMS*, February 2002 <http://www.imsproject.org/question/>
- [3] Questionmark Computing <http://www.questionmark.com>
- [4] Dublin Core specification <http://dublincore.org/>
- [5] Ariadne Educational Metadata <http://ariadne.unil.ch/Metadata/>
- [6] IEEE LOM specification <http://ltsc.ieee.org/doc/>
- [7] Wellington, S J, White S and Davis, H C "Populating the Testbank: Experiences within the Electrical and Electronic Engineering Curriculum" *Proc. 3rd Annual CAA Conference 2001*. Quality Assurance Agency for Higher Education, "Subject Benchmark Statement – Engineering" 2000 <http://www.qaa.ac.uk/cmnetwork/benchmark/engineering.pdf>
- [8] QAA for Higher Education "Handbook for Academic Review" 2000 <http://www.qaa.ac.uk>
- [9] QAA for Higher Education "Code of practice for the assurance of academic quality and standards in higher education. Section 6: Assessment of students" 2000 <http://www.qaa.ac.uk>
- [10] Zakrzewski S and Bull J. "The mass implementation and evaluation of computer-based assessments" *Assessment & evaluation in higher education*. - Vol.23.2 1998 pp. 141-152.
- [11] McKenna, C. and Bull, J. "Designing effective objective test questions: an introductory workshop". *Proc. 3rd Annual CAA Conference 1999* 253-257
- [12] Nielsen, J. "Guerrilla HCI: Using Discount Usability Engineering to Penetrate the Intimidation Barrier" *Designing Web Usability: The Practice of Simplicity* edited by Jakob Nielsen and Robert L. Mack published by John Wiley & Sons, New York, NY, 1994. [http://www.useit.com/papers/guerrilla\\_hci.html](http://www.useit.com/papers/guerrilla_hci.html)
- [13] Draper S.W, Brown M.I., Henderson F.P and McAteer E., "Integrative evaluation: An emerging role for classroom studies of CAL" *Computers & Education* vol. 26.3 1996, pp.17-32.
- [14] Nielsen, J. (1995), *Severity Ratings for Usability Problems*, <http://useit.com/papers/heuristic/severityrating.html>
- [15] White, S. and Davis, H. C. (2001) "Innovating Engineering Assessment: A strategic approach". *Engineering Education: Innovation in Teaching Learning and Assessment Proceeding First Annual IEE Engineering Education Symposium*. 2001.