WHAT'S IN A NAME? - A NEW HIERARCHY FOR QUESTION TYPES

Jane S. Paterson

What's in a Name? - A New Hierarchy for Question Types

Jane S Paterson
Scottish Centre for Research into On-Line Learning and Assessment
School of Mathematics and Computer Science
Heriot-Watt University
Edinburgh

j.s.paterson@hw.ac.uk

Abstract

Any attempt to compare assessment engines is difficult for several reasons. One of these is the terminology that is used to identify question types. As computer assisted assessment develops and extends, new assessment systems are introduced. It is a competitive sector and for those commercial companies involved, a measure of uniqueness is advantageous. All too often this can result in an undue emphasis on finding ways of naming question types to produce the largest number. Close scrutiny reveals that many of these types are derived from the same basic structure with different formatting. The clear cut naming of the initial question types during the first few years of computer assisted assessment worked well but advances in the technology and innovative approaches to assessment are making this convention difficult to sustain.

The work of the IMS QTI group (IMS QTI project 2002) is very valuable and the issue of question types is partly addressed by them. A new structure and naming convention for question types that can be implemented by all interested parties is needed urgently.
There are two aspects to this.

1. A naming convention that would interest those involved in IMS QTI standards and build on the work already undertaken. (the technical sector)
2. A naming convention for the authors, users, academics and researchers interested in what question types are available. (the non-technical sector)

The advantages of such a hierarchy would include

- progress in interoperability
- progress in the use of item banking
- stronger focus on the aims of assessment
- greater awareness of the true question types available

This paper proposes such a hierarchy developed from a non-technical viewpoint but with a sound structure as a basis for discussion, development and to motivate interest in research in this area.

Introduction

The indications are that any new hierarchy will be more complex if it hopes to address the issues that currently cause confusion. Such a hierarchy has to be built on strict criteria to provide question type naming that:

- covers all existing true types of question
- indicates uniqueness
- is capable of expansion in each of the key levels
- is logical, concise and easy to grasp
- provides a framework for the technical sector to develop a standard tagging based on the hierarchy of question naming.

Why have a New Hierarchy?

There are well known assessment engines that claim to have 18 (Question Mark 2002) or even 39 (CIAD 2002) question types or styles. But is this useful or accurate? A question author when comparing question types should be able to identify distinct differences between them and not be left with the dilemma of choosing but wondering why two particular types are available if they operate identically. Similarly there is an urgent need to highlight the differences in question types such as algebraic and numeric which are commonly thought of as one single type. Although most engines will provide numeric question marking there are only a few that can automatically mark algebraic expressions.

A dilemma is reached at the outset when installing a particular assessment system. How can the institution be sure of meeting the needs of their teaching staff without a detailed knowledge of what the assessment systems can provide? Finding the most appropriate engine is therefore time consuming, is unlikely to cover all users and is unlikely to be investigated by a specialist author in each user faculty. A convention in question naming, accepted by all, would at least allow potential users to identify the best system *for their purpose*, saving cost, time and confusion.

What 'Question Types' exist at present?

The answer to this question depends greatly on the assessment engine being considered. Here is one list of question types from the assessment engine CUE (Paterson, 2002).

Judged Mathematical Expression	*Multiple choice*	*Word match*
Hotspot	*Essay*	*Hint*
Multimedia		

The recently published Blueprint (Bull and McKenna, 2001) suggests a general list of question types suitable for computer-assisted assessment. The IMS QTI project also produced a table of question types listed in version 1.2 of their best practice and implementation guide but this focuses on the coding aspect and is perhaps more technical than most users of on-line assessment would like.

It is apparent looking at these three lists that most readers will be able to identify another engine or report with a 'different' question type or name and so the problem multiplies. The question is 'how different are the question types from engine to engine?'.

For example JCloze is a name given to a *fill in the blanks* question type within Hot Potatoes (Half Baked Software, 2002). Another example is: *drop down list* in one engine can be identical to *multiple choice* or even *select a phrase* in others. Can *select a phrase* in fact be treated as a type when it could be presented as a *multiple choice* selection or as a *hotspot?* This is not a criticism of any particular engine but hopefully demonstrates the complex nature of any comparison between assessment engines. Initially the question type names used were well understood, as the community of researchers and developers was compact. With the advances in on-line assessment and a growing interest in interoperability there is a clear need for a revised common understanding of what each question type can do, how it operates and as far as possible, a knowledge of how the question will appear on the screen.

Theoretical Aspects of a New Hierarchy

There have been several very good reports on the requirements of a computer assisted assessment system such as that provided by the Scottish Computer Assisted Assessment Network (SCAAN, 2002). Although this report identifies desirable features in questions, it does not provide information on what is available. However every question author is faced with choices when preparing assessments. From a teaching and learning point of view, question construction has to address the recognised criteria for good assessment (Earl, Land and Wise, 2000). These criteria equally apply to on-line assessment and can affect the choice of question type for a particular task. They have the potential of forming a basis for the new convention. Thus questions have to be:

reliable:
The question should be constructed in such a way that every attempt at it will produce consistent results. Thus a question type must be robust and not open to different interpretations. Random parameters should be chosen carefully.

valid:
The question should test what the student has actually been taught. This may seem obvious, but if part of a learning outcome has been omitted then any assessment on that part is invalid. The question should not introduce extra learning outcomes. A complex structure in a question type could render the assessment invalid if the students have not had some experience of the format in advance. Again any choice of random parameters needs careful consideration.

useful:
To be useful, the assessment must help the student to progress and reinforce their learning. The student has to be aware of what to expect and how the assessment is to be presented. This has close links with ensuring that a question is valid. The students should be familiar with a question construction

or given the chance to practice on similar styles beforehand. It is not useful to the students to produce an all singing-all dancing question type if in fact the students do not learn or reinforce their learning by answering it. Feedback is an important aspect in producing useful questions.

fair:
All students who take the assessment should have an equal chance of scoring full marks. Problems in this area may occur through the wording of a question. Issues on formatting can prejudice question types and render them unfair. The question author should be aware of the problems of accessibility and alternative types should be considered where problems occur. The formatting and choice of buttons should be unambiguous. On-line assessment is far reaching across cultural and ethnic boundaries and this aspect must also be considered in choosing question types.

cost-effective:
Any question type should be justifiable not only in terms of academic content, enhancement to learning and motivation for the students but also in terms of the most efficient question type to produce the required results for the students, the lecturers and the institution as a whole.

A New Basis for the Hierarchy

With these criteria in mind, the common question types take on a different perspective. The focus changes and leads to an emphasis on the following attributes.

What does the question hope to assess? - **OUTCOME**
What input is the student required to take? - **ACTION**
In what manner is this best achieved? - **STYLE**
How is the question best presented? - **FORMAT**

Taking each of these aspects in turn, compatibility with existing question types can be considered.

OUTCOME: This is perhaps the starting point for any author. What does the question hope to test in terms of academic content? It is crucial at this stage for the author to have a precise outcome in mind. Good questions rely on a clearly thought out, well-defined purpose. When this decision has been taken, the next aspect follows naturally. This level is unlikely to be included in the question type naming hierarchy but is the foundation stone for the decisions that follow. It is important however in the construction of item banks.

ACTION: The author is faced with choices, not in what is being tested but on what physical input is required by the student. This is fundamental to the choice of question type. The action selection also influences the skills tested - an aspect that authors need to recognise before making the choice.
At present, this author suggests that there are at most two unique actions:

Selection (using a mouse click)
Data entry (by typing in a string)

The commonly recognised question type names fit into these two categories as follows:

SELECTION: Multiple choice, Hotspot, Drag and drop, Graphing, Mazes
DATA ENTRY: Text entry, Numeric, Algebraic, Gap fill, Crossword

As the reader will no doubt be thinking, there are others or perhaps some of these type names are not recognised - this only serves to reinforce the current state of question type naming. There are of course some assessment systems that allow for several actions within one question and such a question type could be described as *COMPLEX.*

STYLE: This in fact includes the existing types shown in the last paragraph and sub-sectioning them into *selection* or *data entry.* If the decision is taken to use *selection* then it follows that those question types involving some data input by the user are not available. The choice of style can be dictated by the topic being assessed. For example it is unlikely that extended biology questions will be answerable using *selection* action. The hierarchy begins to develop.

FORMAT: This will vary from system to system. It may be dependent on the style chosen or there may be further options. For example, the layout of multiple-choice questions can vary greatly. The type of choice can range from drop down list to simple button choice; the layout can be single or multiple column choice; the impression can be chosen using coloured coordinated text, background and font options.

The Naming Convention and Coding Possibilities

Any acronyms and coding can cause frustration among those who are not familiar with them. However it is unlikely that any new naming convention can continue using the simple but once effective method of describing the question type by what it can do. Such a hierarchy has to be built on strict criteria to provide question type naming that:

- covers all existing true types of question
- indicates uniqueness
- is capable of expansion in each of the key levels
- is logical, concise and easy to grasp
- provides a framework for the technical sector to develop a standard tagging based on the hierarchy of question naming.

There are several options possible:
1. continue but control the naming convention as it stands
2. set up a numeric coding system
3. set up an algebraic coding system
4. use an IP address convention but using letters

1. To continue with the naming system as it stands has its attractions. At a basic level there is a common understanding of what type of question will be

generated as, for example, a hotspot. However, innovative use of existing question types has removed the clear-cut differences between these base types and there are more attempts to incorporate unique features of one type into others. To adhere to the new criteria may lead to a very cumbersome description of each question type. There is no underlying structure to provide for new names and the issues of uniqueness cannot be addressed easily using this convention.

2. A numeric code would be appealing to those with a mathematical or science background and to those involved in IMS standards. It lends itself well to use within computing systems but is likely to be harder to understand and unlikely to be popular with the non-scientific sector. Whilst it could be structured to provide a strong underlying basis, to indicate uniqueness and to be capable of expansion it would be meaningless to the layperson or those trying to gain a foothold in computer based assessment.

3. An algebraic coding suggests a system similar to using acronyms. Although there is a great deal of opposition to using acronyms, they are an efficient and quick way of remembering and referring to rather lengthy titles. In general, each letter in an acronym stands for one word that makes it difficult to use for this system. For example, it is not easy to use only one letter for the style attribute, as provision is needed for both 'essay' and for 'equation'.

4. Extending the ideas of an algebraic code further leads to a possible solution: in IP address format there are four groups of three digit numbers separated by a dot. If a similar idea is adopted using letters it will give four groups of three letters separated by dots. This fits in well with the four proposed attributes of outcome, action, style and format.

This paper proposes a system based on this style. The proposal however does not suggest that such question types will be discussed using a 12-letter code. The code may be very useful for inclusion in standards and in the storage within question banks but even a cursory glance at a possible structure will demonstrate that it is possible to formalise the naming in a coherent and understandable way as follows:

- Allow 'z' to be used as a blank space thus shortening the bulk of the names but allowing for expansion
- Place the attributes in the four groups in the order 'action', 'style', 'format' and 'outcome'. This allows for the outcome to be identified when necessary but may be dispensed with in general terms.
- In conversation use the *action* letters, the *style* name and the *format* letters. An example follows later.

Demonstration of the convention

***ACTION**:* At present only four actions are possible.
SZZ, TZZ, NZZ, MZZ - Select, Type, No Input or Mixed

***STYLE**:* For each of the first codes a further three letters are added to select the styles available within that action. Examples follow.

SZZ -

MCZ - multiple choice (one choice)
MRZ - multiple response (more than one choice)
HSZ- hotspot

TZZ -

TMZ - text match (word or phrase)
ALZ - algebraic (markable not display)
EQZ - equation (markable)

FORMAT*:* This attribute may well require all three letters and will depend on the choices made under *action* and *style*. Examples follow.

SZZ-

MCZ -

BFR - Button, four column, random order
RLF - Radio button, listed, fixed order

SZZ-

HSZ -

SIM - Single graphic, Image, Multiple correct

Thus a common type may be logged as SZZ-HSZ-STM-AEZ meaning a question type where the answer is **S**elected on a **H**ot**S**pot with **S**ingle graphic of **T**ext and a **M**ultiple correct answers to test the outcome skill of **A**pplication at an **E**asy level.

No-one is ever likely to talk in terms of such a code as it stands but by calling this question type an 'S - hotspot - STM', most listeners could identify precisely the type of question likely to be provided and be confident in making a decision on its suitability for their needs.

The Way Forward

Compilation of a full version of the code is underway. SCROLLA (2002) is currently working on the worldwide evaluation of assessment engines and this research will inform the progress on this hierarchy. However any new convention will not be successful unless it has the support of all those who use it and work with it: this includes researchers, academics, tutors, lecturers, developers and not least the commercial suppliers of assessment engines. SCROLLA welcomes collaborative research on this new hierarchy from any interested parties.

References

Bull, J. and McKenna, C. (2001) *Blueprint for Computer-assisted Assessment*. Luton: CAA Centre

CIAD, (Centre for Interactive Assessment Development) (2002) *TRIADS*, Derby: University of Derby
http://www.derby.ac.uk/ciad/ciastyles.html (13 May 2002)

Earl, S. Land, R. and Wise, J. (2000) *Assessment, Evaluation and Support*. MSc Computer-enhanced Mathematics Education, Edinburgh: Napier University

Half Baked Software, (2002) *Hot Potatoes.* Canada: University of Victoria
http://web.uvic.ca/hrd/halfbaked/ (13 May 2002).

IMS QTI Project, (2002) *Question And Test Interoperability Specification.* IMS Global Learning Consortium
http:www.imsproject.org/question/qtiv1p2pd/imsqti_asi_bestpdv1p2.html#1278750 (13 May 2002).

Question Mark Computing, (2002) *Perception v3.* London
http://www.questionmark.com (13 May 2002)
http://www.scrolla.hw.ac.uk/reportlist.html (13 May 2002)

Paterson, J. S. (2002) *The CUE Assessment System*. MSOR Connection **2(**2) Birmingham: LTSN
http://ltsn.mathstore.ac.uk/articles/maths-caa-series (13 May 2002)

SCAAN, (Scottish Computer Assisted Assessment Network) (2002)
http://www.scaan.ac.uk/public_docs/academic-requirements.doc
(13 May 2002)