Training staff to use a new MCQ Generation methodology (CAREGen) in the UK Electricity Distribution Industry

Robert M. Foster
University of Wolverhampton, UK

Abstract

The first part of this article presents the results from a recent survey which demonstrates increased usage of Multiple Choice Question (MCQ)-based assessments within a UK company. The survey begins with a trend analysis of 5 consecutive annual totals of KACEs (Knowledge Acquisition Confirmation Events) arising from the company's use of MCQs. This is followed by a comparison of the proportions of KACEs that occurred in formative, summative and refresher assessment contexts. An alternative method for categorizing KACEs is then applied to the data from selected years to show how MCQ use has expanded into new content sub-domains. The conclusion from the survey is that MCQ-Creation and delivery are becoming increasingly important in the effective communication of corporate knowledge, rules and procedures. The recommendation is that investment is required in the development of staff skills in MCQ-Creation and the associated task of writing effective approved documents from which MCQs are derived.

Since one of the report's recommendations was that more staff should receive MCQ-Creation training, a detailed design for a MCQ-Creation workshop is provided in the second part of this article. The same workshop design was applied during the delivery of a MCQ-Creation workshop at a recent education conference when the presenter was awarded the 'Best workshop Award'. The article concludes with a call for an online version of the MCQ-Creation workshop to be created.

1. Introduction

UK Companies are required by legislation such as Health and Safety at Work, etc Act 1974 [1] to provide information, instruction, training and supervision to ensure the health and safety at work of their employees. Some UK companies have introduced systems of training and assessment that use Multiple Choice Questions (MCQs) as part of their response to these requirements. This paper begins by presenting the results from a survey of the application of MCQs within a UK company which illustrates how the quantity and scope of MCQ usage in addressing this requirement are increasing [2].

Figure 1 provides a clear illustration of how MCQs are becoming increasingly important in the effective communication of corporate knowledge, rules and procedures. The measure used in the survey is a 'KACE' (Knowledge Acquisition Confirmation Event) which simply counts the number of times a user has clicked the correct response to a Multiple Choice Question during an assessment.

The survey begins by looking more closely at the figures behind the above graph which provide a trend analysis of MCQ usage over 5 consecutive years. There follows an analysis of the proportions of MCQ usage in 2012 in each of three assessment categories: formative [3], summative [4] and refresher [5] assessments. Then there is an alternative presentation of the totals for selected years which show how this company is using MCQs to address assessment needs in an increasingly diverse range of content sub-domains. The conclusion from the survey is that MCQ-Creation and MCQ-Delivery are becoming increasingly important in the effective communication of corporate knowledge, rules and procedures in this company.

An early response from the team responsible for MCQ-Creation was to investigate the possibility of using software to generate MCQ test items automatically [6], [7], [8]. However the conclusion of this investigation is that it is likely to be several years before systems of this kind can produce outputs that could be used in this company.

In the meantime, the MCQ-Creation team has applied meta-cognitive analysis techniques [9] to examine the manual process of MCQ-Creation. Alongside the resolution of some concerns among training specialists, other outputs from this meta-cognitive analysis include a recommendation concerning the most appropriate format of MCQ test item [10] and a specification for a formal MCQ-Creation methodology[11]. There were also some ideas for a MCQ-Creation workshop. These ideas were applied during the delivery of a MCQ-Creation workshop at a recent conference after which the presenter was awarded the 'Best Workshop' award.

After presenting the survey, this article provides a description of the aims and structure of the MCQ-Creation workshop. Only a brief summary of the MCQ-Creation methodology is included, since a detailed description is available elsewhere[11]. However, the article does include ideas for

simplifying the methodology and for enhancing the constructivist learning dialogue between the readers and writers of company approved documents. The recommendation is that the MCQ-Creation workshop should be delivered to staff who are responsible for MCQ-Creation, and also to those who are responsible for preparing the approved documents which define the foundation knowledge for MCQ-Creation.

2. Background

The UK Health and Safety at work, etc Act 1974 [1] specifies that every employee has a duty of care for the health and safety of themselves and their colleagues:

Health and Safety at Work, etc Act 1974 – Section 2

> *"2. (1) It shall be the duty of every employer to ensure, so far as is reasonably practicable, the health, safety and welfare at work of all his employees.*
>
> *(2) Without prejudice to the generality of an employer's duty under the preceding subsection, the matters to which that duty extends include in particular:*
>
> *(A)... (B) ... (C) the provision of such information, instruction, training and supervision as is necessary to ensure, so far as is reasonably practicable, the health and safety at work of his employees; (D)... (E)...."*

Companies often provide the instruction, training and supervision required by the Health and Safety at Work, etc. Act 1974 through formative[3], summative[4] and refresher[5] knowledge check assessments so that knowledge gaps can be identified and then addressed.

Formative assessments [3] can be carried out during training to allow trainees to build new knowledge into their world view at a pace that suits them. Summative assessments [4] can be used immediately after training to confirm that trainees have retained the important knowledge from their training in short term memory, and Refresher assessments [5] can be used to confirm that trainees have retained the knowledge from their training in long term memory. In this way a coordinated system of formative, summative and refresher assessments can be very effective in supporting the activity of managing corporate knowledge.

In recent years the method for delivering assessments that provide the instruction, training and supervision required by the Health and Safety at Work, etc., Act 1974 has changed. In place of paper based questionnaires, managers now use Computer Based Tests (CBTs) consisting of Multiple Choice Questions (MCQs). An assessment context in which this change has been specified in the company's approved document library is illustrated below:

ST: OS7D – Relating to Audit Procedures for Operational Authorisations – Paragraph 3.1

> *"All Senior Authorised and Authorised Persons who hold an authorisation for HV Operational Work (11SW, 33SW, 66SW, 132SW and restricted variations) shall complete an annual CBT test to the satisfaction of an Examining Officer qualified to examine for that authorisation."*

An important motivation for moving towards systems that include automatic marking has been the significant increases in staff numbers in recent years (in 2011 the existing staff numbers of 2,200 were increased to over 6,000). However the move towards using MCQs within CBTs is also supported by the case study quoted in a recent experiment [5] which in addition to demonstrating how MCQs used in refresher contexts can be evaluated, also described how careful application of a new MCQ-Creation methodology in response to a specifically identified learning need, provided a cost effective method for filling a significant knowledge mis-match between two categories of staff. It has also been noted that if staff who are seeking to gain increased levels of authorization know that they will need to pass a CBT as part of the authorization interview then they will revise their learning materials more vigorously than if the CBT had not been included in the authorization process.

3. Survey Method and Results

The survey presented in this paper makes use of the 'KACE' as a measure of MCQ usage. A 'KACE' has been defined [5] as a Knowledge Acquisition Confirmation Event. A 'KACE' is considered to have occurred each time a user of a Computer Based Test (CBT) clicks the correct response button within a Multiple Choice Question (MCQ).

The first graph in the survey (Figure 1) and the associated data table (Table 1) shows a count of all KACEs that occurred within the company within the specified years.

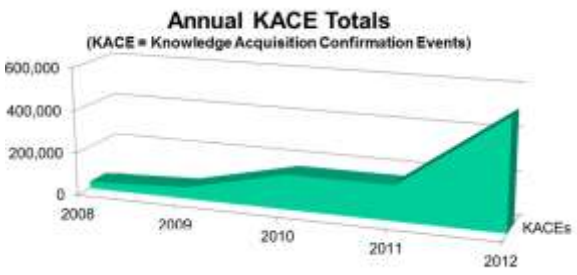

Figure 1. KACE totals between 2008 and 2012

Table 1 does not include KACE counts from 1990 – 2007 even though CBTs consisting of MCQs

have been used in the company since 1990. The reason for this omission is that annual KACE totals between 1990 and 2007 did not change significantly. The interesting changes are seen to have occurred since 2008.

Table 1. KACE totals between 2008 and 2012

Year	KACE total Note: 1 x KACE = 1 x correct response to a MCQ test item by a staff member
2008	32,890
2009	53.655
2010	162,680
2011	164.031
2012	511,602

In Figure 2 and the associated data table (Table 2) the KACEs for 2012 are presented in categories according to the assessment type for which the MCQ was created and used.

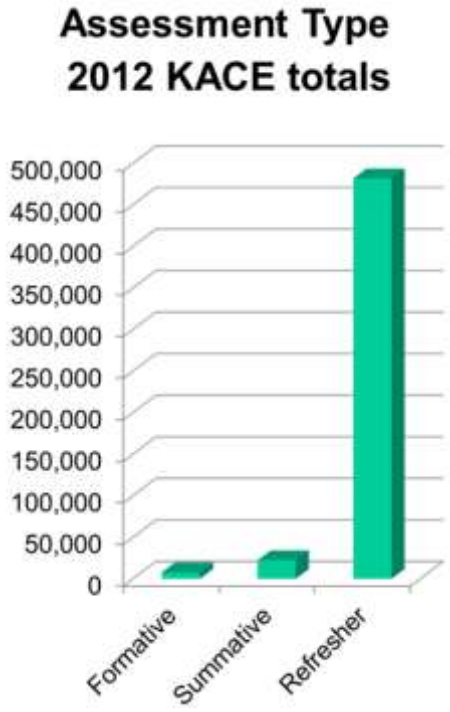

Figure 2. 2012 KACE counts by assessment type

This table only includes KACE counts for 2012 since only very low KACE counts for formative and summative assessment CBTs were found in years prior to 2012.

Table 2. 2012 KACE counts by assessment type

Assessment Type	2012 KACEs Note: 1 x KACE = 1 x correct response to a MCQ test item by a staff member
Formative	7,800
Summative	22,149
Refresher	481,653

The figure associated with Table 3 (Figure 3) provides illustrations of some of the sub-domains covered by these MCQs and attempts to illustrate how MCQs facilitate the dialogue about the company's policy documents:

Figure 3. MCQs facilitate dialogue in a range of sub-domains

In Table 3 the KACEs for 2008, 2010 and 2012 are categorised according to the content sub-domain which is covered by the MCQ in which the KACE occurred:

Table 3. KACE counts in 2008, 2010 and 2012 by Content Sub-domain

Content sub-domain	2008 KACEs	2010 KACEs	2012 KACEs
Chainsaws	0	0	28,840
Heavy Plant	7,620	25,380	31,380
Operational Safety	0	47,480	271,685
Safety Health and Environment	25,270	55,920	104,160
Specific Risk	0	31,500	45.588
Apprentice Training	0	2,400	29,949

MCQ response data that includes counts of KACEs continues to be gathered, but the results presented in this report are those collected before 31st December 2012.

4. Discussion of survey findings

Perhaps the most surprising table among the three presented in this survey is Table 2 which indicates a very low level of MCQ usage for the widely researched assessment contexts of Formative[3] and Summative[4] assessment. The text accompanying the table states that "only very low KACE counts for formative and summative assessment CBTs were found in years prior to 2012". It is also interesting to note that the number of sub-domains for which Computer Based Tests have been produced has increased during the surveyed period as indicated by Table 3.

The most obvious message from this survey is provided by the general trend in Table 1. Usage of MCQs in the featured company is clearly increasing. That table also shows that there have been step changes in MCQ usage within the featured company between 2009 and 2010 and again between 2011 and 2012. The rise in MCQ usage in 2010 is likely to have been caused by the introduction of CBT assessments covering many more content sub-domains, as is shown in Table 3. The same explanation cannot be given for the 2012 rise since Table 3 shows that nearly all content areas that existed in 2012 had already been introduced in 2010. Closer examination of Table 3 shows that the increases apply in all content areas, with the most significant proportional increases appearing in the areas of Apprentice Training assessments and Operational Safety Training assessments.

Table 2 indicates that the vast majority of KACEs occurring in years prior to 2012 have been in the context of a Refresher assessments. If we combine this observation with the observed step change in KACE counts in 2012 and the observation in the Background section that there were significant staff number increases in 2011 then the fact that the step change in KACE counts occurred a year after the new staff arrived is consistent with the observation that most KACEs occur during refresher assessments.

It is clear that effective processes for creating, delivering and maintaining MCQs are going to become increasingly important in the effective communication of corporate knowledge, rules and procedures in this company. We therefore consider next the processes that have been observed to be applied during MCQ-Creation.

5. Analysis of MCQ-Creation processes

Following the realization that MCQ-Generation software was unlikely to deliver a short term solution to the challenges facing the MCQ-Creation team, they initiated a meta-cognitive analysis of the manual processes of MCQ-Creation that were being applied. The study involved some of the company's training specialists and during their interviews they raised the concern that an increasing usage of MCQs might reduce the level of constructivist learning dialogue within the company. Their point was that such dialogues would previously have occurred during face to face training courses and that increased use of MCQs might reduce the level of such dialogue within the company.

However, on further investigation, it became clear that downward pressure upon costs had already caused managers to send fewer staff on face to face courses. Also, local Examining Officers, who are responsible for awarding and refreshing authorization certificates and conducting field checks, reported that using MCQs as part of their interviews had actually led to an increase in the level of constructivist learning dialogue within their team.

In addition to resolving the concerns from the training staff, the study also generated evidence based recommendations for the most appropriate format of MCQ test item and a specification for a formal MCQ-Creation methodology.

The steps of the methodology are presented below in Table 4.

Table 4. CAREGen Methodology

Step	Description
1	Define the objective of the MCQ routine using a CSLO
2	Identify the most appropriate source documents
3	Explicate (and if necessary Add) Coherence Relations for sentences that meet certain criteria and then re-work them into CRST-compliant CSLOs
4	Extract candidate antonym pairs for each of the identified sentences
5	Apply construal operations in the context of identified antonym pairs
6	Generate MAC test items by inserting generated components into a template.

The meta-cognitive analysis also produced some ideas for a MCQ-Creation workshop and when these ideas were applied during at a recent conference, the presenter was awarded the 'Best Workshop' award. The next section contains a more detailed description of the workshop including the promotional material

and the timetable. The recommendations section contains some ideas for customizing the workshop for different categories of attendee.

6. MCQ-Creation workshop

The title and tag line for this workshop use the acronym 'MCQ'.

6.1. MCQ-Creation Methodologies Workshop

"At MCQ-Creation we discuss empirical studies of MCQ creation methodologies and then suggest improvements."

The term 'MCQ' does not require further explanation for those who are interested in this topic and in any case, the logo contains the expansion of the term.

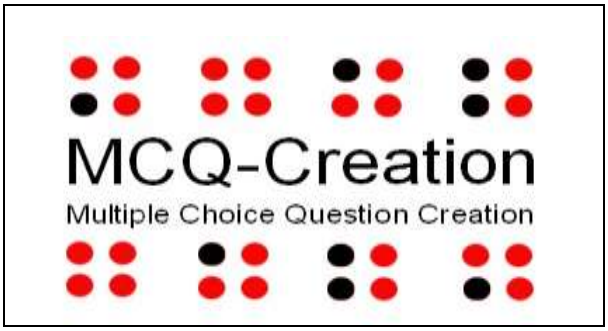

Figure 4. Logo for the MCQ-Creation workshop

The objectives of the workshop are spelled out in the call for papers:

"The specific objectives of assessment authorities and the languages of learner populations are continually changing. This implies that assessment tools, which are used to confirm accurate transfer of knowledge to learners within a domain of discourse, must also change.

The MCQ-Creation Workshop brings together educationalists from industry, governmental examining bodies, universities and schools to examine the merits and pitfalls in traditional processes for creating Multiple Choice Question (MCQ) test items. The output from the workshop will be proposals for new (or adapted) MCQ creation methodologies that are appropriate to the domains defined by the presenters.

The submissions process, important dates and submission template are presented on the workshop website. Submissions are emailed to a workshop-specific email address (MCQ-Creation@testcentres.co.uk).

Details for submissions are clearly defined and the date for submissions is deliberately set for after the workshop. This allows delegates to learn from and then apply the steps of the methodology and then report on their success to future sessions of the workshop, during the case study session.

Presenters at MCQ-Creation will give an overview of their domain of discourse (ie the context of their assessments) and will describe how they have either:

- *Described and evaluated an established formal method for MCQ test item creation as applied within their domain of discourse*

  or

- *Described and evaluated a demonstrably NEW MCQ test item creation methodology within their domain of discourse. "*

The call concludes with a welcome message which also aims to create the relaxed atmosphere that is necessary for effective learning:

"Welcome to the MCQ Creation Methodologies Workshop in conjunction with the LICE (London International Conference on Education). I look forward to hearing your ideas ... and sharing some of mine of course !"

A summary of the content of the workshop that was delivered at LICE 2012 is provided in the following table:

Table 5. MCQ-Creation Workshop

MCQ-Creation Methodologies Workshop
Welcome and introductions
Aims of the MCQ-Creation workshop
Lecture 1: Defining knowledge domains *(includes an introduction to my knowledge domain by summary of my research)*
Exercise 1: Define YOUR knowledge domain
Lecture 2: What should be tested / measured?
Exercise 2: What will you test / measure in YOUR knowledge domain?
Lecture 3: Validity
Lecture 4: How do we measure achievement?
Exercise 3: What are the available types of MCQ?
Lecture 5: The MCQ Creation Methodology

Exercise 4: Important components of the MCQ-Creation
Lecture 6: MCQ Creation Guidelines
Exercise 5: Important Guidelines for your knowledge domain
Case Studies: - High Voltage Cable Jointing - Heavy Plant Lifting - Apprentice Progress Monitoring
Exercise 6: What have you learned?

7. Conclusions

Clearly MCQ-Creation and Delivery are becoming increasingly important in the effective communication of corporate knowledge, rules and procedures in this company. KACE counts are increasing in all categories of MCQ assessments, although extending the use of MCQs into Formative [3] and Summative [4] assessment contexts might lead to further benefits. This recommendation is supported further by recent research [10], which has shown that when compared to traditional assessment methods, CBTs consisting of MCQs in the MAC (Multiple Alternate Choice) format deliver more comprehensive feedback within formative assessments and more targeted identification of knowledge gaps during summative assessments. Other research [12] has shown how the process of MCQ-Creation can be more closely linked to the approved documents upon which training materials are based and can therefore provide an effective method for delivering formative and summative assessments of staff knowledge following training in current corporate knowledge.

The meta-cognitive analysis of the MCQ-Creation processes at this company has revealed many of the properties of a healthy constructivist learning dialogue. Perhaps the creation and delivery of a MCQ-Creation workshop would produce several benefits in addition to the maintenance of the banks of MCQ test items which are clearly needed. Simplification of some of the steps through the application of customized software might lead to further improvements in the quality of the items, as the designers are released to consider the educational as well as the assessment qualities of the test items they produce.

8. Recommendations

8.1. Recommendations from the survey

The survey presented in this paper shows how the number of KACEs made possible by the delivery of corporate knowledge using MCQs has risen significantly each year since 2008. This annual rise provides supporting evidence for the assertion that if MCQs are to be used to deliver, reinforce and refresh corporate learning as part of a system for managing corporate knowledge, then greater investment is needed in the development of staff skills in MCQ-Creation and the associated task of writing effective approved documents from which MCQs are derived.

In addition to the general rise in MCQ usage, some usage patterns have also been identified which suggest that increased application of MCQs to formative and summative assessment contexts might prove beneficial. Perhaps some kind of 'conversion' of MCQ test items that have proved useful in refresher assessments might be possible to make them suitable for use in formative and summative assessment contexts.

8.2. Recommendations for MCQ-Creation

In response to the strong case made by this survey and associated meta-analysis of manual MCQ-Creation techniques, the main recommendation is that subject experts who are likely to get involved in MCQ-Creation, should attend a MCQ-Creation workshop. Perhaps the cost of attending the workshops could be reduced through a customization of the methodology so that it blends in more easily with the existing MCQ-Creation processes.

Another recommendation is that in addition to subject experts who are likely to be involved in creating MCQs, such as trainers and staff development officers, there might also be a benefit for subject experts who write the company's approved documents attending MCQ-Creation workshops as well. Again, a customised version of the workshop might be appropriate which focuses upon the issues of policy writers. There are likely to be significant benefits if the writers of policy documents understand the MCQ-Creation challenges that will be faced by those who must implement the rules and procedures they define in their documents.

If the language and style of documents that define the company's approved rules and procedures facilitated manual MCQ-Creation, then there is also a chance there would be improved results from MCQ-Generation software solutions. For example the language definition might include a lexicon in which each term had only one accepted definition. Also, there might be rules that imposed a limit upon the number of words in each sentence, and this can make a significant difference to execution time for software that uses generative grammar models.

Case studies in which the above recommendations have been implemented would make very interesting submissions to the next MCQ-Creation workshop. It would be particularly interesting to hear of customized versions of the workshop leading to new coverage of new content sub-domains within an organisation or increasing the levels of attendance that have been approved by managers.

9. References

[1] UK Government, (1974), 'Health and Safety at Work, etc Act' (http://www.hse.gov.uk/legislation/hswa.htm) (Access date: 15 May 2012)

[2] Foster, R.M. (2013) – 'Improve MCQ-Creation Skills To Support Corporate Learning' - Ireland International Conference on Education 2013.

[3] Crooks, T., (2001), "The Validity of Formative Assessments". British Educational Research Association Annual Conference, University of Leeds, September 13-15 2001

[4] Glickman, C.D., Gordon, S.P. & Ross-Gordon, J.M. (2009) – Summative Assessment

[5] Foster, R.M., (2012) – 'Evaluating an application of the CAREGen MCQ Creation Methodology'–London International Conference on Education 2012

[6] Mitkov, R., and L. A. Ha. 2003. "Computer-Aided Generation of Multiple-Choice Tests." In Proceedings of HLT-NAACL 2003 Workshop on Building Educational Applications Using Natural Language Processing, pp. 17-22. Edmonton, Canada.

[7] Mitkov, R., L. A. Ha, and N. Karamanis. 2006. "A computer-aided environment for generating multiple-choice test items." Natural Language Engineering 12(2): 177-194.

[8] Brown J.C., Frishkoff G.A. Eskenazi M., 2005 "Automatic Question Generation for Vocabulary Assessment" Processing (HLT/EMNLP), pages 819–826, Vancouver, October 2005. © 2005 Association for Computational Linguistics

[9] Degrave, W.S. et al., (1996) "PROBLEM-BASED LEARNING - COGNITIVE AND METACOGNITIVE PROCESSES DURING PROBLEM ANALYSIS", Instructional science, 24(5), 1996, pp. 321-341

[10] Foster, R.M., (2010) – 'Multiple Alternative Choice test items (MACs) deliver more comprehensive assessment information than traditional 4-option MC test items ' – London International Conference on Education 2010

[11] Foster, R.M. (2012a) "Using a new MCQ Generation methodology (CAREGen) in the UK Electricity Distribution Industry" In 'International Journal of Digital Society (IJDS), Volume 3, Issues 1 and 2, March/June 2012 643 - 651

[12] Foster, R.M., (2011) 'Creating a High Voltage Cable-Jointing knowledge check using the CARE generation methodology' – London International Conference on Education 2011.

