How Do You Like Your Learning? E, M or C?

Rebecca M^{c}Keown
Cranfield University
Cranfield, Milton Keynes, MK43 0AL, UK
r.mckeown@cranfield.ac.uk

Dr John Huddlestone
Cranfield University
Cranfield, Milton Keynes, MK43 0AL, UK
j.huddlestone@cranfield.ac.uk

ABSTRACT

The British Army's move towards blended learning delivery (a combination of classroom and technology-based learning) for Military Annual Training Tests (MATTs) instigated this research in order to evaluate the effectiveness of e-learning and m-learning in delivering MATTs training content, alongside extant classroom-based training. An evaluation of user reaction was assessed by way of questionnaires and workshops. Evaluation of Learning effectiveness was addressed using an empirical controlled experimental design. A pre and post-test assessment allowed comparison of the pre and post-training scores to determine the effect of each training delivery method. The sample group comprised soldiers and officers from the Field and Territorial Army and cohorts of soldier and officer recruits. A total of 425 participants were recruited for the research and randomly allocated to one of the three methods.

The user reaction towards the concept of training using newer technologies was encouraging; the m-learning package was considered the most popular. Participants reacted positively to having a choice of different types of learning content, either "*mixing things up bit"* or working to their preferred modalities of learning. Flexibility in terms of when and where learning was carried out was seen to be of benefit. The results from the learning effectiveness evaluation revealed that that there were some improvements between pre and post-test scores, and some of those improvements reached statistical significance. The newer technologies (e-learning and m-learning) were as effective as classroom delivery.

It was recommended that further development of e-learning and m-learning packages needs to be user-centred and embrace instructional design principles. Consideration also needs to be given to using e-learning and m-learning as part of a blended solution to delivering MATTs training, alongside instructor led classroom sessions – particularly for emotive-based topics. It is also recommended that research on knowledge retention of materials delivered by e-learning or m-learning be explored.

ABOUT THE AUTHORS

Rebecca M^{c}Keown joined the Department of Integrated Systems at Cranfield University as a Research Fellow in 2004. She is currently engaged in a number of military training research projects. In addition to her military work, Rebecca carries out research for the commercial aviation industry. She also works with the career development team for students on the Cranfield University MBA programme. She holds a BSc (Hons) in Psychology from the Open University and a Masters Degree in Occupational Psychology from Cranfield University.

Dr John Huddlestone is a Senior Research Fellow in the Department of Integrated Systems at Cranfield University. He leads military human factors research being conducted under the auspices of the Departments work within the Human Defence Capability Science and Technology Centre. His research interests include simulation, training and task analysis. Prior to joining the University he was a commissioned officer in the Royal Air Force. As a training specialist, he gained extensive experience in training design, aviation training development, flight simulation and the development of computer based training. He holds a PhD in Applied Psychology from Cranfield University, an MSc in Computing from Imperial College London and a BEd in Physics from Nottingham Trent University. He is a Member of the British Computer Society and is a Chartered IT Professional.

How Do You Like Your Learning? E, M or C?

Rebecca M[c]Keown
Cranfield University
Cranfield, Milton Keynes, MK43 0AL, UK
r.mckeown@cranfield.ac.uk

Dr John Huddlestone
Cranfield University
Cranfield, Milton Keynes, MK43 0AL, UK
j.huddlestone@cranfield.ac.uk

INTRODUCTION

The British Army are moving towards blended learning delivery, a combination of classroom and technology-based learning, for Military Annual Training Tests (MATTs). These tests are a series of training packages covering various mandatory training topics including Personal Weapons Training, Battlefield Casualty Drills, Values and Standards, Operational Law and Counter Improvised Explosive Devices.

MATTs have to be completed on an annual basis by serving British Army personnel and are currently delivered via classroom sessions using scripted PowerPoint presentations and training aids such as DVD clips and scenarios. The purpose of the move towards blended learning is to provide flexibility in training delivery.

E-learning has already become a staple training method; advances in technology have meant that mobile devices are now capable of delivering high quality information via software applications, commonly referred to as 'Apps', and the concept of m-learning has evolved. In defining m-learning, Degani et al (2010) highlighted that the nature of learning is more diverse than that of e-learning *"Any activity that allows individuals to be more productive when consuming, interacting with, or creating information, mediated through a compact digital portable device that the individual carries on a regular basis, has reliable connectivity, and fits in a pocket or purse"* (p8). This definition also highlights the more personal nature of m-learning as mobile devices tend to play a large part in both the business and the social lives of many.

The use of Apps sparked interest across the British Army and it was felt that the benefits may enable MATTs training material to become more accessible thereby facilitating revision and completion to improve first time pass rates and reduce skill fade. It was also felt that the inclusion of Apps in a blended learning package may increase motivation to learn and realise efficiencies through course compression.

The E-learning and m-learning training materials developed for the British Army each included a range of media options (modes) to deliver the training materials. The e-learning packages included text to read (in the form of a pdf file), DVDs, Scenarios (read or video clips), Interactive content and self-assessment. The m-learning content also included text to read, interactive content and self-assessment content and, in addition to these modes, also included material presented by way of games and podcasts (audio material). Although the content varied slightly between the different modes of delivery, the same learning objectives were addressed in order to provide a comparative learning experience across e-learning, m-learning and classroom-based training.

With the decision to move towards a blended learning package made, and the training materials developed, research was instigated with the aim of evaluating the effectiveness of e-learning and m-learning in delivering MATTs training content, alongside extant classroom-based training.

Training Evaluation

Kirkpatrick (1998) proposed a training evaluation framework which was deemed suitable for this research as it provides an established structure for evaluation and is embedded within the Defence Systems Approach to Training. Kirkpatrick proposed that there are four levels at which training can be evaluated:

- Reaction: What did the trainee think and feel about the training?
- Learning: Did the trainee actually learn anything?
- Behaviour: did the learning change the trainee's subsequent behaviour in the workplace?
- Results: Did the training change in the trainees' behaviour to achieve the desired effect?).

This research addressed the 'Reaction' and 'Learning' elements of Kirkpatrick's framework. These elements of evaluation are closely interlinked; positive reactions

to training are considered to be a prerequisite for training to be effective. Positive reactions to training do not constitute sufficient evidence that the training has been effective therefore it is essential to determine that learning has actually taken place.

The elements identified for inclusion within the 'Reaction' component of this research were derived from the Human-Computer Interaction domain and related to usability and the user experience. Measuring these elements is necessary in order to obtain a complete understanding of users' needs and improve the product to provide a better user experience. Bevan (2008) defines usability as *"acceptable effectiveness, efficiency and satisfaction"* (p13). Zhang and Adipat (2005) identified nine factors that are most often evaluated in m-learning contexts – 'User Satisfaction', 'Learnability' (how easily users can complete a task first time), 'Efficiency' (how fast users can complete a task), 'User Error', 'Effectiveness' (completeness and accuracy in achieving goals), 'Simplicity' (degree of comfort in using the product) and 'Comprehensibility' (how easily can the content be understood). de Sa, Carriço and Duarte (2008) suggested evaluation of more practical issues such as dexterity of using the device should also be considered.

In order to assess learning, Kirkpatrick and Kirkpatrick (2006) proposed that a control group should be used, if practical, to enable comparisons to be made as to the effectiveness of the new training intervention. They also proposed that an evaluation of knowledge, skills and/or attitudes both before and after a training programme should be carried out:

- Use a test related to the content of the training event to measure knowledge.
- Use a performance test to measure skills.
- A questionnaire survey can be used to measure attitudes.

Therefore the objectives of this research were to:

1) Understand how the learning experience was viewed.
2) Discover whether learning occurred.

METHOD

In order to fit in with research timescales and participant availability, two components from two of the MATTs were selected by the British Army for inclusion in the trials. MATT6 Equality & Diversity (ED) - the 'Respect for Others' sub-component of 'Values and Standards' training which underpins the moral component of fighting power. MATT7 Law of Armed Conflict (LOAC) - a sub-component of 'Operational Law' which trains all ranks of the British Army in accordance with national and international legal obligations'. These particular components were chosen as they were representative of different types of content. MATT6 (ED) covers general compliance issues whereas MATT7 (LOAC) covers military only content.

The e-learning materials were developed for delivery via CD-ROM and the m-learning materials for delivery via iPod Touch™ devices from one of two sources; either as a stand-alone App downloaded directly onto the device or from a Learning Management System (LMS) - an App downloaded onto the device but linked to a database provided by the supplier and automatically updated when the device is connected to the internet. The e-learning and m-learning packages comprised a variety of modes of delivery e.g. Case studies or scenarios (read or video clips), audio, interactive content such as assessment in the form of quizzes and games, and reading material.

In order to understand how the learning experience was viewed, user reactions were surveyed using a mix of questionnaires and workshop discussions. A user reaction questionnaire, containing 15 items with a Likert-style response format rating from 1 (strongly disagree) to 5 (strongly agree), was designed to address the elements of usability and user reaction identified in the literature review. Sections for qualitative feedback were also included. The workshop discussions were based on the output of a Product Reaction Card exercise. This methodology was developed by Benedek and Miner (2002) as a way of measuring intangible aspects of the user experience quickly and easily. Product reaction cards are a set of word cards containing both positive and negative words and phrases to cover a wide variety of dimensions. These cards form the basis for a sorting exercise and discussion about the product under evaluation. Benedick & Miner's original product reaction card list of words comprised 118 words which was too lengthy to be completed in the time available for the research so a re-development process was undertaken by members of a working group in the field of mobile learning and training. The most common words identified as being pertinent to the research were collated and the original list was revised to contain 23 words (13 x positive, 10 x negative).

In order to investigate the learning effect of e-learning, m-learning and classroom training, an empirical controlled experimental design was used. A pre and post-test assessment design was employed to allow

comparison of the post training scores against pre-training scores to determine the effect of each training method. A parallel form of pre and post-trial MATTs knowledge assessments were developed for this purpose. There were two independent variables in the research study:

- IV1 - the method of training delivery (e-learning, m-learning or classroom)
- IV2 - the time at which knowledge was assessed i.e. pre-training or post-training. This variable was tested via repeated measures, with each participant participating in both conditions.

To investigate whether learning occurs, the total number of correctly answered pre and post test scores were collected and used as the dependent variable.

The design allowed a comparison between pre and post-test knowledge assessment scores for each of the three methods of delivery to determine if there were significant differences between pre and post-test scores; and whether the methods of delivery had an effect on learning.

Sample

The participant groups considered to provide a comprehensive cross-section of the potential training audience were:

1) An 'ab initio' cohort comprising sub-groups of both soldier and officer recruits. This group was selected for inclusion in the trials as there was not available data on skill fade for the refresher training groups. By including an ab initio group, any learning effects were more likely to be identified.
2) A 'refresher training' cohort comprising combat and technical support functions of the Field Army, combat and technical support functions of the Territorial Army (TA) and an Officer sub-group.

A total of 425 participants were recruited for the research and randomly allocated to one of the three methods of training delivery.

Data Analysis

To analyse the qualitative 'Reaction' data a Template Analysis approach (developed by King, 1994) was used with the product reaction card words providing the template for analysis of the data from the workshops. Data coding gave structure to the data and enabled a final template of thematic trends to be identified. This final template was used to analyse the content the collated qualitative data from the user reaction questionnaires.

The quantitative data from the user reaction questionnaires was analysed statistically. To produce an overall 'desirability' score, the ratings for all delivery methods were totalled, taking into account the reversal of negatively-orientated questions and the means were calculated.

In order to establish whether learning had occurred, and whether these differences were due to the method of delivery, the total number of correct answers from the pre and post-test knowledge assessments was recorded for each ab-initio participant and mean scores were calculated. To test for significance with any differences found, the results were analysed via a mixed (between-within) Analysis of Variance (ANOVA). Where appropriate either post-hoc analyses, in the form of Tukey's Honestly Significant Differences or and paired t-tests were carried out to offer greater explanatory power to the results.

RESULTS

Objective 1) How is the learning experience viewed?

This section presents the results from the quantitative element of the user reaction questionnaire. The qualitative results are presented within the discussion in the next section.

The e-learning and for m-learning groups were asked to specify which modes of learning they used most during their training session, and which was their favourite. The results for the e-learning group showed that, for MATT6 (ED), 70% of participants used the interactive quiz most frequently and 47% deemed it to be their favourite. The reading material was less frequently used (10% of the participants) with 14% rating this mode as their favourite. The case studies were least frequently used (5% of participants) and least popular; only 9% rating this mode as their favourite. For MATT7 (LOAC), the interactive modes (games and quiz-style assessments) were used most (54% of participants) and also most popular as 51% rated these as being their favourite. The scenario modes were used less (27% of participants) and rated as favourite by 30% of respondents. The least used and least popular of the modes was the reading material which was used most frequently by 8% of participants and rated as favourite by 7%.

The results for the m-learning group showed that, for MATT6 (ED), 49% of participants used the interactive

quiz most frequently and 43% deemed it to be their favourite. The game was most frequently used by 20% of participants with 18% rating it as their favourite. The reading material and audio modes were least used at 10% and 9% of participants respectively and also rated least favourite (8% and 10% respectively). In the MATT7 (LOAC) training materials, the interactive and quiz-style assessment modes were also the most frequently used (56% of participants) and the most popular, with 53% rating these modes as favourite. Once again, the reading material and audio modes were least used at 13% and 7% of participants respectively and also rated least favourite (9% and 8% respectively).

An examination of the mean desirability scores showed that the m-learning method received the highest ratings for both MATT6 (ED) and MATT7 (LOAC). See Tables One and Two below:

Table 1: Mean Desirability Scores for MATT6 (ED) E-learning, M-learning and Classroom Learning

		N	Mean	SD
MATT6 (ED)	e-learning	109	43.54	8.907
	m-learning	183	44.52	9.320
	classroom learning	88	44.08	9.212

Table 2: Mean Desirability Scores for MATT7 (LOAC) E-learning, M-learning and Classroom Learning

		N	Mean	SD
MATT7 (LOAC)	e-learning	119	44.82	8.853
	m-learning	165	45.61	9.174
	classroom learning	86	42.93	8.025

Objective 2) Did Learning Occur?

Tests were carried out to see whether group differences were statistically significant. In order to do so, the data were entered into a mixed (between-within) ANOVA with post-hoc analyses in the form of Tukey's Honestly Significant Difference and/or paired t-test where appropriate. The results are presented below:

Ab Initio Sample Group – MATT6 (ED)
The results of the mixed (between-within) ANOVA showed that post-test scores were higher than pre-test assessment scores for the ab initio sampling group ($F(1,120) = 11.238$, $p<0.05$). The officer trainee pre and post-test scores were significantly higher than those of the soldier trainees ($F(1,120) = 23.98$, $p<0.05$) which could reasonably be expected.

There was also a significant interaction effect between time and the method of delivery ($F_{(2,120)} = 24.94$, $p<0.05$). In order to discover the nature of this effect a series of paired t-tests were carried out. Where the statistical significance of the paired t-tests was $p\leq0.05$ the figures are shown alongside the mean pre and post-test assessment scores in Table 3.

This data confirmed that the significant differences between pre and post-test scores were found in the e-learning and classroom learning groups, for the soldier trainee sampling group. This means that e-learning and classroom-based learning were more effective for training MATT6 (ED) for this particular group of participants.

Ab-initio Participant Group - MATT7 (LOAC)
The results of the mixed (between-within) ANOVA again showed that post-test scores were higher than pre-test assessment scores for the ab initio sampling group ($F_{(1,120)} = 48.88$, $p<0.05$). As with MATT6 (ED) the officer trainee pre and post-test scores were significantly higher than those of the soldier trainees ($F(1,120) = 121.13$, $p<0.05$) which could reasonably be expected. A paired t-test paired t-test data revealed that significant differences between pre and post-test scores were found across all three methods of learning (e-learning, m-learning and classroom learning group) and were found with both the soldier and officer trainee sampling sub-groups. Where the statistical significance of the paired t-tests was $p\leq0.05$ the figures are shown alongside the mean pre and post-test assessment scores in Table 4.

This means that e-learning m-learning and classroom-based learning were equally effective effective for training MATT7 (LOAC) training content for the ab initio sampling group as a whole.

Refresher Participant Group - MATT6 (ED)
The results from the ANOVA showed there was a significant difference in pre and post-test assessment scores ($F_{(1,269)} = 4.543$, $p<0.05$) and a significant main effect of sampling group on assessment scores ($F(1,269) = 4.695$, $p<0.05$) indicated there were differences in terms of pre and post-test scores between the three sampling groups. Tukey's Honestly Significant Difference post-hoc analysis of the results revealed that there were significant differences between Officers and Regular soldiers and between TA soldiers and Regular soldiers, but not between Officers and the TA. There were no significant differences due to different methods of learning.

To provide additional explanatory power, a series of paired t-tests were carried out. Where the statistical significance of the paired t-tests was $p \leq 0.05$ the figures are shown alongside the mean pre and post-test assessment scores in Table 5.

This paired t-test data revealed that a significant improvement in learning only occurred for Regular soldiers undertaking classroom-based training.

Refresher Participant Group – MATT7 (LOAC)

The results from the ANOVA showed there was a significant difference in pre and post-test assessment scores ($F_{(1,226)} = 13.639$, $p<0.05$) and a significant interaction effect between time and sampling group ($F(2,226) = 3.222$, $p<0.05$) indicated there were differences in terms of pre and post-test scores between the three sampling groups. Tukey's Honestly Significant Difference post-hoc analysis of the results revealed that there were no significant differences between the sampling groups (i.e. there were no significant differences between the Officer, Regular and TA groups), and that there were no significant differences between the methods of learning. To investigate further a series of paired t-tests were carried out. Where the statistical significance of the paired t-tests was $p \leq 0.05$ the figures are shown alongside the mean pre and post-test assessment scores in Table 6.

The paired t-test data has revealed that e-learning and classroom-based training were most effective for the Officer sub-group, e- learning was most effective for the Regular soldier sub-group and m-learning most effective for the TA soldier sub-group.

In summary, the results from the ab-initio cohorts showed an improvement between pre and post test scores on MATT6 (ED) for all but one m-learning group, with some differences reaching statistical significance. Of more practical interest, statistically significant differences were found between pre and post-test scores for all of the MATT7 (LOAC) conditions. Whilst there were statistically significant differences between pre and post-test scores for the Refresher group, there were no significant differences between the different methods of delivery. The practical significance of these findings is small. We concluded that there were some improvements between pre and post-test scores, and some of those improvements reached statistical significance. The newer technologies (e-learning and m-learning) were as effective as classroom delivery.

DISCUSSION

Objective 1) How is the learning experience viewed?

The data to meet this objective were captured by means of User Reaction Questionnaires and workshops and this section provides discussion of the key themes identified as well as a discussion of the results presented above.

The quantitative results from the User Reaction Questionnaire showed the m-learning method of delivery was considered most popular. This is due to the novelty of this method of learning as evidenced by the following. It was very interesting to note that participants had far more to say or to contribute to discussions on the utility of m-learning than they did to discussions on the other training methods. It is probable that learners have become very used to having a PC or laptop, either at work or at home, and more accustomed to e-learning. In contrast, m-learning using the iPod Touch™ devices was seen as "*novel", "modern",* and *"slick and cool"* by participants. 73% of ab-initio participants and 75% of refresher participants reported they owned smart phones (either an iPhone™ or similar). The take-up of smart phone technology has been rapid, with users motivated to explore the seemingly never ending applications of this new technology.

The topic of how to interface with a PC or laptop did not arise in any of the workshop discussions, suggesting that learners were broadly conversant with computer use and interacting with computers. As for interfacing with the iPod Touch™ devices, participants were generally unanimous in their view that there was no real need to explain how to use the devices and interact with the content, and that even novice users would be able to "*pick it up fairly quickly".* A small number of workshop participants branded themselves as novices, but did report being able to easily use the devices to undertake the requisite learning packages.

Being able to interface with a device, whether it be a PC or smart phone, is clearly an essential pre-requisite for all learners, as the user needs to be able to interact and engage with the learning content. Users from both the e-learning and m-learning groups generally found the MATTs learning packages easy to use, but felt they were not very intuitive, as they did not allow users to take control over how they navigated through the learning material. Workshop participants from both groups also spoke of frustration with not being able to navigate through lengthy video, audio or reading sections – with a strong call to be able to search through this type of content. Others talked about the frustration

of not being aware of how far through a learning component they actually were, and not knowing how much additional time was needed to complete the task.

Workshop participants from the e-learning and m-learning groups also called out for more interaction. In both cases they were either trying to obtain more information, definitions to support something they had just seen, to be able to ask questions or to discuss the learning with others. The e-learning participants suggested that pop-up text boxes could be used to provide immediate access to information where and when it is needed, and the m-learning workshop participants suggested the use of a forum for discussion and questions (e.g. a web-based forum). In addition to the workshop participants generating some useful development suggestions, it appears that learners are thinking about taking some responsibility for their learning. They are becoming active, rather than passive learners. If the learning packages are developed properly, then e-learning and m-learning have the potential to further encourage active learning.

In terms of interaction, the m-learning workshop participants also suggested using video clips, punctuated with questions such as "*What would you do now?*" followed by immediate feedback which they can compare with their proposed course of action. As noted in the results section, it was also suggested that in developing scenario-based learning, much can be learned from commercial 'Off-The-Shelf' games, where branching storylines are used, and users learn from experiencing the outcome of decisions taken, in a benign context.

As noted above, feedback is an essential component of the interaction between the user/learner and the software, but it is also key to user engagement. Providing accurate and timely feedback was seen as being important to both e-learning and m-learning users. By virtue of the nature of the content developed for the research, it was the e-learning workshop participants who were most vocal about poor feedback from the training materials. They reported that the MATT7 (LOAC) quiz-style assessment only provided feedback that their answer was incorrect and no indication of what the correct answer was. In addition to this, on completion of the test, they were scored as having failed even if they had scored a fairly high number of correct answers. A lack of feedback, or inappropriate feedback, can impact on an individual's motivation to progress through remaining material, result in frustration and a loss of user engagement.

It is very clear from the above discussion that there is a need to engage with end users throughout the design and development process. It is important in terms of helping to design and develop training content, understanding how users interact and engage with the content, as well as de-glitching 'bugs' in the system; a number of which were identified from the data obtained.

The suitability of e-learning and m-learning as platforms to deliver MATT6 (ED) and MATT7 (LOAC) content was an important discussion topic. Doubts about the suitability of e-learning and m-learning to support MATT6 (ED) as a stand-alone delivery method came from participants' experiences of the games modes of delivery. The e-learning game came under criticism as it could not distinguish between opinion and intention. The m-learning game was criticised for being pitched at too low a level and trivialising the subject matter. In both instances participants felt the need to engage in some discussion based around realistic scenarios in a classroom setting.

Doubts about the suitability of e-learning and m-learning to deliver MATT7 (LOAC), were less focused on the training content per se, but on the perceived importance of the subject matter, due in no small part to the consequences of getting it wrong. In this case, participants suggested that MATT7 (LOAC) should be delivered by a suitably qualified training officer, who would be able to put the learning material into context, explain the reasoning behind particular courses of action, and be able to answer questions. What was distilled from the workshop discussion is that both MATT6 (ED) and MATT7 (LOAC) comprise some emotive content, which requires interpretation, and some factual content. It was felt that whilst factual content lends itself to e-learning and m-learning delivery, facilitated discussion was needed to deal with some of the issues and tensions associated with emotive content to put it onto context and embed the learning. The story that emerged was that participants saw utility in both MATT6 (ED) and MATT7 (LOAC) e-learning and m-learning, but as part of a blended solution alongside classroom-based training.

Alternative uses were envisaged by the m-learning workshop participants, such as using the materials for revision, or the mobile device as a portable source of reference rather than for training per se. Participants also saw benefits in being able to be flexible in terms of when and where they carried out their learning. Some of the e-learning participants considered conducting their training in the office at work, or at home in the evening although it should be noted that the TA participants were more minded to the latter option. M-learning participants saw opportunities to train while on a long bus journey, while being transported to a firing

range, and during short periods of down time. Participants did make it clear that whilst they may choose to undertake MATTS training using e-learning or m-learning delivery during their free time, it should not be expected and time needs to be allocated during the working day for MATTs training. On a similar note, participants also expressed that whilst they may own a smart phone that can support MATTs learning content, they should not feel pressured to load the content onto their personal devices or to purchase devices to support the learning content. Participants were more relaxed about loading content onto PCs or laptops – largely because they were not seen as being as personal as a mobile device.

For the most part, participants reacted positively to having a choice of different types of learning content, either "*mixing things up bit"* (certainly more than can be achieved with classroom training) or working to their preferred modalities of learning. Not all participants agreed, and some thought the content was dull and boring. For the purpose of the trial, participants explored all of the different learning modes and, as noted from the user reaction questionnaire results, participants largely made most frequent use of the quiz-style assessment and game-based Apps or the interactive content with reading material and audio material being the least frequently used. The results for 'favourite' component broadly followed the results for most frequently used.

Learners will have preferences for different learning modalities (e.g. visual, auditory and kinaesthetic), and it is important that in delivering training as many modalities are supported as possible. It is important that learners are not tempted to only choose a limited range of components, and in so doing restrict the range of the learning content that they access. This is imperative when the different learning options cover different content, as in the case of MATT 6 (ED) and MATT7 (LOAC). A number of participants mentioned that they would be tempted to go straight to the quiz-style content and use the outcome of that to determine what other content they review. Others suggested that they may just take a quick look at the material, or only look at the components that they think they would be engaged with. As a result, there are real concerns that if learners have the opportunity to decide how they interact with the content that, unlike structured classroom-based training, some learning points and learning content may be missed.

Whilst the empirical trial data supports that learning has occurred, albeit that the differences in pre and post test scores only reached statistical significance in a proportion of the trials (predominantly with MATT7 (LOAC)), and no differences in test score could be attributed to any one particular learning method, in general, participants from both m-learning and e-learning cohorts reported positively towards undertaking MATTs training using these newer technologies. However, there are technical bugs and glitches to be sorted out and there is a need to engage with end users as modifications are made to the learning material, followed by further empirical testing. Consideration also needs to be given to using e-learning and m-learning as part of a blended solution to delivering MATTs training, alongside instructor-led classroom sessions.

Objective 2) Does learning occur?

Ab-Initio Participant Group

Evidence from the MATT7 (LOAC) ab-initio participants showed that there were significant differences between pre and post-test scores after undertaking MATTs training. This indicates that the e-learning and m-learning training methods were as effective as classroom training. The results from MATT6 (ED) were less clear; the Soldier m-learning sub-group had higher pre-test scores than those obtained post-test. Further examination highlighted that the m-learning pre-test score was also higher than the pre-test scores for the other Soldier sub-groups. The start-state of the m-learning sub-group was no different to that of the e-learning and m-learning cohorts and a thorough interrogation of the data did not yield an explanation for this anomaly. There were also differences between the pre and post-test scores for the Officer sub-group but these did not reach statistical significance.

The difference in results between MATT6 (ED) and MATT7 (LOAC) could be due to the nature of the content. With MATT6 (ED), a number of questions could realistically be answered based on a popular understanding of equality and diversity issues and common sense. Answering the questions for MATT7 (LOAC) required exposure to the training material, as it was novel to all but a few of the officer cohort, arguably making for a more robust measure of the utility of the three methods of learning. In practice, there were no significant differences between the methods of delivery.

Refresher Participant Group

Evidence from the refresher participants undertaking MATT6 (ED) training has shown that there were no significant differences between pre and post-test scores across any of the three modes of learning, with the exception of Regular soldiers undertaking classroom learning, for whom the difference achieved statistical significance. The situation was slightly different with

MATT7 (LOAC) in as much as there were statistically significant differences between pre and post-test scores for the TA m-learning sub-group, Officers e-learning and classroom sub-groups, and Regulars classroom sub-group. There were no significant differences due to the mode of delivery.

It has to be noted, however, that the average pre-test scores for each MATT (averaged across all modes of learning and all the refresher sampling sub-groups: Officer, Regular and TA) were already at a fairly high level. The mean for MATT6 (ED) pre score was 12.96 (SD 1.509) marked out of 14, and for MATT7 (LOAC) the pre score was 15.74 (SD 1.780) marked out of 17, so the margin for improving knowledge was always going to be slim for the refresher group. It was for this reason that the ab-initio groups were included in the study, as they had little or no prior exposure to the MATTs training content before the trial. For purposes of comparison, the mean MATT6 (ED) pre score for the ab-initios was 10.77 (SD 3.372) and for MATT7 (LOAC), the mean pre score was 9.95 (SD 4.268).

The analysis of the refresher participants according to recency of completion of MATTs training did not reveal any significant differences between pre and post-trial scores for MATT6 (ED), but significant differences were found for MATT6 (ED) for participants who had never undertaken MATT7 (LOAC) training and those who had last undertaken MATT7 (LOAC) training over 12 months prior to the trials. The fact that the significant differences occurred with MATT7 (LOAC) rather than with MATT6 (ED) was not wholly unexpected given the discussion about the nature of the content e.g. factual vs emotive subjects. Of practical significance is that the data suggests the refresher interval for MATT6 (ED) could be extended beyond 12 months and possibly beyond 12 months for MATT7 (LOAC), although further research would need to be conducted to confirm this.

CONCLUSIONS

In general, participants from both m-learning and e-learning groups responded positively towards the concept of undertaking MATTs training using these newer technologies. Some limitations in terms of the suitability of the e-learning and m-learning platforms to deliver MATTs content were reported; factual content was seen as being appropriate for e-learning and m-learning delivery, but emotive content, or elements that require interpretation are better taught in a classroom with the opportunity to ask questions and develop discussion.

A number of constructive comments were made by participants relating to their interaction and engagement with the e-learning and m-learning packages, but their comments did reveal a number of shortcomings, including several technical 'glitches' which need to be remedied. Another issue to be considered was that of how to encourage all learners to engage with all of the e-learning and m-learning materials, and not to be selective and restrict the range of material they engage with.

Participants welcomed the flexibility and portability offered by e-learning and m-learning, including the options to study in their own time, but expressed the need for time to be made available during the working day to complete the mandatory MATTs training packages.

RECOMMENDATIONS

Further development of the MATT6 (ED) and MATT7 (LOAC) e-learning and m-learning training packages needs to be more user-centred, i.e. engagement needs to be made with end-users during the continuation of the development process. Any training material developed needs to remedy the identified software issues and embrace instructional design principles, especially in terms of feedback and navigation.

Consideration also needs to be given to using e-learning and m-learning as part of a blended solution to delivering MATTs training, alongside instructor led classroom sessions.

Further research could include a more fine-grained analysis of the influence of training recency on the differences between pre and post-trial assessment scores, to help determine whether refresher intervals can be extended. It is also recommended that research on knowledge retention of materials delivered by e-learning or m-learning be explored.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the contribution of Dr David Swift and Major Graham Yare (DTrg(A)). We would also like to express our gratitude to all of the service people who participated in the planning and executing of this research project at DTrg(A), DETS(A), RTMC, AFC, ITG-HQ, RMAS.

Table 3: **MATT6 Mean Pre and Post-trial Scores for Ab-initio Participants**

Method of learning	Sampling group	N	Pre-Trial		Post-Trial		
			Mean score	Standard deviation	Mean Score	Standard deviation	Sig $p \leq 0.05$
M-Learning	Soldier	29	11.31	2.47	11.14	2.83	-
	Officer	14	12.36	2.47	13.00	1.47	-
E-Learning	Soldier	32	9.50	3.88	11.59	2.01	0.010
	Officer	21	12.67	1.20	13.19	1.17	-
Classroom	Soldier	29	9.45	4.06	12.38	1.86	0.003

Table 4: MATT7 Mean Pre and Post-trial Scores for Ab-initio Participants

		N	Mean score (pre-trial)	Standard deviation (pre-trial)	Mean Score (post-trial)	Standard deviation (post-trial)	Sig $p \leq 0.05$
M-Learning	AFC	29	8.97	3.00	10.69	3.47	0.018
	RMAS	14	13.93	2.76	16.57	3.18	0.001
E-Learning	AFC	32	7.66	2.81	11.81	3.54	0.000
	RMAS	21	15.19	1.57	16.71	1.23	0.001
Classroom	AFC	29	7.38	3.44	9.62	3.22	0.004

Table 5: MATT6 Mean Pre and Post-trial Scores for Refresher Participants

Method of learning	Sampling group	N	Mean score (pre-trial)	Standard deviation (pre-trial)	Mean Score (post-trial)	Standard deviation (post-trial)	Sig $p \leq 0.05$
M-Learning	Officers	44	13.16	1.18	13.34	0.96	-
	Regular	41	13.24	1.07	13.07	1.21	-
	TA	45	13.22	1.00	13.22	1.22	-
E-Learning	Officer	24	13.04	1.52	13.29	0.86	-
	Regulars	23	12.70	1.69	13.13	1.42	-
	TA	20	13.00	1.17	13.20	1.24	-
Classroom	Officers	23	13.43	0.73	13.30	1.06	-
	Regulars	37	12.03	2.57	12.89	1.58	0.024
	TA	21	12.90	1.14	13.48	0.60	-

Table 6: MATT7 Mean Pre and Post-trial Scores for Refresher Participants

Method of learning	Sampling group	N	Mean score (pre-trial)	Standard deviation (pre-trial)	Mean Score (post-trial)	Standard deviation (post-trial)	Sig $p \leq 0.05$
M-Learning	Officers	21	15.67	1.39	16.43	1.47	-
	Regulars	17	15.82	1.63	16.12	1.54	-
	TA	52	16.00	1.88	16.67	1.71	0.010
E-Learning	Officers	21	15.67	2.13	17.05	1.20	0.007
	Regulars	23	16.26	1.18	16.52	1.70	-
	TA	19	15.47	1.98	15.37	4.22	-
Classroom	Officers	29	15.62	1.78	16.72	1.03	0.002
	Regulars	31	15.32	1.85	15.97	1.58	0.043
	TA	22	15.82	1.65	15.82	2.44	-

REFERENCES

Benedek, J. & Miner, T. (2002). 'Measuring Desirability: New Methods for Evaluating Desirability in a Usability Lab Setting'. Proceedings of UPA 2002 Conference, Orlando, Florida, July 8 – 12 2002.

Bevan, N. (2008). Classifying and Selecting UX and Usability Measures. Proceedings of the International Workshop on Valid Useful User Experience Measurement, Reykjavic, Iceland.

DeGani, A., Martin, G., Stead, G. And Wade, F. (2010). Mobile Learning Shareable Content Object Reference Model (m-SCORM) Limitations and Challenges (N09-35). Office of Naval Research.

de Sa, M., Carriço, L. And Duarte, L. (2008). A Framework for Mobile Evaluation. CHI 2008, Florence, Italy pp 2673 – 2678, ACM Press, April 2008 cited in Proceedings of the International Workshop on Valid Useful User Experience Measurement, Reykjavic, Iceland.

King, N. (1994). Template Analysis. In Symon, G. and Cassell, C. (Eds), (1998). Qualitative Methods in Organizational Research: A Practical Guide. London: Sage.

Kirkpatrick, D. L. (1998). Evaluating Training Programs: The Four Levels. (2nd Ed). San Francisco: Berrett-Koehler.

Kirkpatrick, D. L. And Kirkpatrick, J. D. (2006). Evaluating Training Programs: The Four Levels. (3nd Ed). San Francisco: Berrett-Koehler.

Zhang, D. and Adipat, B. (2005) Challenges, Methodologies and Issues in Usability Testing of Mobile Applications. International Journal of Human-Computer Interaction, 18, 293 – 308.