# The e-assessment burger: supporting the before and after in e-assessment systems

Journal Article

# The E-assessment burger: Supporting the Before and After in E-Assessment Systems

Anne Adams [1], Gill Clough[1]

[1] Open University, Walton Hall, Milton Keynes,
MK17 8TJ, UK
gillclough@open.ac.uk

**Abstract.** This paper describes a threshold concept-driven e-assessment system that supports teachers in writing effective formative multiple-choice questions, creating quizzes tailored to students' learning pathways. The system, which has been co-designed with teachers, acts as the 'bun' on either side of an 'e-assessment burger' pedagogically scaffolding quiz creation (the top of the bun), integrating the quiz within personalized learning trajectories (the burger) and feeding the results back to the learners and teachers to guide the direction of future learning pathways (the bottom of the bun). The evaluation with 26 students in 3 subjects across two schools identified that supporting the before and after e-assessment empowers a shift in teachers' encouragement for student ownership of assessment, guiding their learning pathways. Teachers also provide insights into how the system scaffolding and visualisations inspired changes to sequencing learning and teaching practices. In conclusion the changing role of assessment within a school ecosystem is debated.

**Keywords:** threshold concepts, e-assessment, learning pathways, formative assessment, visualisations, secondary education, diagnostic

## 1 Introduction

E-assessment can transform learning by empowering the learner with systems that provide personalised, adaptive computer marked and interactive feedback that is both formative and summative [1, 2]. Assessment systems are often focused upon how they technically support, personalise and mark students learning [3]. Automated e-assessment systems have advanced the area of formative assessment processes by providing students with an avenue to learn through the assessment system. Providing automated feedback, confidence based marking and computer adaptive tests adjusting to students responses to previous questions [1, 4, 5]. However, the full potential of e-assessment systems are continually under-used in practice [6]. This could be because often e-assessment developments focus more on technical innovations than how innovations are framed within learning pathways, ownership, agency and organisational eco-systems.

Whilst great advancements in general student-directed learning support have been made there is a slower development in supporting assessment systems' pedagogical underpinning, social/organizational test creation and implementation procedures.

Similarly there needs to be a greater understanding of how e-assessment results are integrated into innovations in student directed learning pathways. We suggest that this is akin to an assessment burger where most research has focused on the middle section of test creation, storage and marking. What is often missing are the links between the e-assessment system and the conceptual gaps in understanding prior pre-existing the assessment, and the link through to inform teaching practices after the assessment, in particular with regard to student-directed learning.

Many of these innovations can be enabled by teachers to empower student directed learning through approaches like flipped learning [7] before and after e-assessment results are received. Within flipped learning students learn the core material through lecturers, online material at home. Students use lessons to explore and test their understanding with the expertise of the teachers to identify if they've 'got it' or not. Assessment can provide a useful trigger for flipped learning identifying what core learning needs to be done by the student and what they need to review further with the teacher. However, whilst many evaluations have identified the benefits of flipped classrooms [8, 7] both in secondary and higher education the interlink with assessment is poorly documented. This presents an example of how innovative assessment systems relates to broad social and organisational teaching and learning approaches. There are, however, more specific pedagogies that can and should be used to provide leverage for challenging and changing agency in assessment procedures.

### 1.1 Background

It has been argued that central to designing assessment systems is their alignment with teaching and assessment methods [9, 10]. Assessment and learning design are therefore clearly interlinked. With this in mind, a key objective for learning has been noted as identifying pathways to deep as opposed to surface learning [11, 12, 13]. Assessment systems should therefore support assessing a student's deep as opposed to surface understanding.

A long history of research has identified that some types of learning have long-term effects fostering generalised deep thinking and the power to transform the individual [14]. It could be argued that a key connection point between assessment and learning processes is that it could provide insights into the students' internalisation of concepts in order to effectively complete the assessment. However, assessment tools are only a proxy for understanding just as a driving test is only a proxy for determining a good driver. The quality of the questions within the assessment tool is important point for assessing a student's depth of understanding. Deep learning and internalisation of concepts is difficult but a necessary objective of learning to transform a students understand rather than simply supporting repetitive mimicry of an understanding through regurgitating information. To support effective questions creation we need to identify pedagogies that review understanding and learning processes towards deeper understanding.

Threshold concepts[1] [15, 16] have become a focal point for understanding conceptual barriers learners encounter towards a deeper understanding of a concept. Their research has pointed towards TCs as a starting point for transformative learning [16]. Threshold Concepts were originally identified in two founding papers by [15, 16] as a 'portal' to a different way of thinking through internalization of concepts without which the learner finds it difficult to progress [16]. They are said to be more than just "key" or "core" concepts [17, 18]. The barriers presented by Threshold Concepts can be so great, they may cause students to fail or give up a subject altogether and research has highlighted the need to focus on effective methods for teaching Threshold Concepts [19]. Although not without their critics, [20, 21], Threshold Concepts have widespread support within the teaching and academic community. Yet there is poor evidence of how they have been applied to assessment systems.

Multiple choice questions has been noted as historically created by teachers and through a 'drill and practice' fashion facilitating a surface learning approach [22]. However, [23] highlight that even the simplest automated question and answer system can support students in understanding topics. This then introduces a key decision within e-assessment of what and when to automate. A key distinction, we would argue, is not the level of complexity of the automated system but it pedagogical underpinning related to the teaching and learning ecosystem within which it is place.

It is often proposed that computer systems should automate repetitive tasks whilst allowing academics to use their expertise and skills within complex assessment procedures [24]. Whilst automation has been reviewed in-depth around the issue of marking procedures, the complex issue of supporting question creation autonomy and shared agency has been less clearly supported. An assessment overview paper by [25] documents one future direction for e-assessment systems as that of supporting students creating tests and reviewing peers tests. There have been some developments in this area on automating the transfer of question items, tests and results data between systems providing assessment authoring tools, item banks and outputs [26, 27]. Nevertheless, the focus has tended to be upon the technologies not the teachers, learners and the learning process. Jordan [2] and Whitelock and Brasher [3] present accounts and reviews on initial research around developing answer-matching rules to support increasing the quality of question creation for question banks. However, whilst invaluable, these reviews do not address the socio-organisational issues of how to use computer-assisted support in constructing questions that provide a better fit with learning design processes.

Bacigalupo et al [6] identified 8 obstacles to the uptake of e-assessment systems, within a specific HE environment, only half of which were technical. Yet, the one identified obstacle, of quality in questions writing, was noted as requiring examples and guidance. Unfortunately the authors answer to this obstacle was to provide training and a guidance manual. It could be argued that supporting teachers creating valid questions is too difficult for technical automation. This then emphasises the issue of how we can empower students to effectively generate questions when they would understandably require more support than teachers. Although this maybe a difficult issue to technically address it does not detract from the value for learning of student created test questions.

[1] Threshold concepts are later defined as 'tricky topics' based upon teachers feedback this related to their practice better than what they considered was a formalised academic term which was a threshold concept in itself.

et. al. Denny et al [22] noted the benefit of student generated questions systems on the students learning as well as pragmatically on limited teacher resources.

The first barrier identified by Denny et al [22] in technically supporting students in creating their own test questions was in question classification, in particular the 'topic' for question creation. The authors dealt with this problem pragmatically rather than pedagogically by using the chapters of the course book as the topics. It was argued that this related strongly to Schulte and Bennedsen [28] post-hoc evaluation study of computer science topics, again driven by subjective practice approaches (i.e. asking teachers to rank the relevance of what they teach) rather than any approaches to learning design or learning models. Luxton-Reilly et al [29] have further developed support for student generated question banks. However, again their focus was upon the systems storage and retrieval rather than any pedagogical guidance for the learning. For example, the guidance tool focused on prompts for clear language and instructions, grammar and formatting.

This paper details the evaluation of a question and answer system that provides a pedagogically underpinned guidance process for question creation and resulting visualisations. The first step in this process is developed by educators to facilitate appropriate depth and guidance in learning. The e-assessment results are similarly connected back to an effective threshold concept focused learning pathway. This provides the foundations for effective questions and answer creation with results that empower students to direct and control their flipped learning experiences.

### 1.2 Aims and Objectives

As has been highlighted by the literature review many e-assessment systems have had a background pedagogically based in the learning outcomes approach to assessment such as the 'drill and practice' approach of many multiple choice systems. We argue that it may not be multiple choice structures that are in itself limited. It could be how these system are designed to fit with the learning process as a simple test of knowledge acquired. This then traditionally places such systems at the end of learning to test progression. However, even if they are placed earlier on in the process they again can provide a blunt 'how much' do you know assessment. The rationale for our approach to assessment is that it should be tied to the learning processes. We aim to identify if this provides a more effective support system for learning development and students ownership of that learning. We review three issues in developing and evaluating this approach and eassessment system:

- Problems for / barriers to learning,
- depth of learning (deep and surface)
- reflection and learning ownership

We seek to address these within this systems through 1) identifying conceptual barriers within threshold concepts (tricky topics) 2) assessment of understanding at a deep or surface level through stumbling blocks associated to quiz quesitons 3) support for reflection through quiz visualisations supporting student reflections and their ownership of learning.

### 1.3 System & Schools Description

The e-assessment tool described in this paper was developed through a series of co-design workshops with teachers and students over an 18 month period.

*The school contexts*

Whilst the learning processes and systems have been used in schools and HE across Europe (i.e. Portugal, Spain, Sweden and Germany). Two UK city schools were involved in these assessment evaluation trials. These schools have a spread of abilities and many of the students in both schools are second language learners. Teachers involved in the trials range in their years of experience and abilities. The pressures on timetables are prevalent and as such the value of systems must be obvious to teachers before they support engagement. Collaborative system design supported this engagement and value for the subsequent system.

*The system collaborative design process*

The project team used a variety of methods and tools to ensure collaborative design, summarised below in Table 1.

Table 1: System Development workshops

| Date | ID | Participants | Methods |
|---|---|---|---|
| 14/11/12 | UKRW1 | 3 biology teachers<br>3 chemistry teachers<br>1 physics teacher<br>1 technology teacher | • Concept mapping<br>• Participant observation notes (Livescribe pen recordings)<br>• Audio recording |
| 14/11/12 | UKTUW 1 | Course Chair<br>2 lecturers | • Concept mapping<br>• Video recording<br>• Audio recording<br>• Participant observation notes (livescribe pen recordings) |
| 28/11/12 | UKRW2 | 2 biology teachers<br>2 biology students<br>3 chemistry teachers<br>2 chemistry students<br>1 physics teacher<br>2 physics students<br>Facilitated by 2 university physics lecturers, 2 JxL researchers | • Student notes (livescribe pen recordings)<br>• Video recordings<br>• Audio recordings<br>• Participant observation notes (livescribe pen recordings)<br>• Focus group debrief audio recordings (teachers) |
| 25/03/14 | UKRW3 | 1 head of technology<br>1 chemistry teacher<br>1 head of drama<br>2 Physics lecturers/ STEM ambassadors | • Video recordings<br>• Eye-tracking<br>• Participant observation notes (livescribe pen recordings) |
| 22/4/14 | OUDQ1 | 1 chemistry teacher | • Participant observation notes (livescribe pen recordings) |

The Diagnostic Quiz is the starting-point for students taking part in the JuxtaLearn Process. The Diagnostic Quiz initiates the process by identifying the key gaps in the students' understanding of the threshold concept, referred to as a Tricky Topic, previously identified by the teacher.

There were two developmental stages of the 'tricky topic tool' developed. The first was based upon a 'word-press' system which acted as a rapid prototyping tool for teachers presenting the process. It was through this process of rapid development that the need for support in question creation based upon the tricky topic problems was identified. The second stage of development moved the 'tricky topic tool' into an integrated 'juxtalearn' system called 'ClipIt' supporting the whole learning process from assessment to creative video making, commenting and learning analytics toolkits. This system was based upon the social networking platform Elgg allowing for integration of different external applications adapted via the appropriate ClipIt API.

*The top of the e-asssessment burger bun*

In the first stage, effectively, the top of the e-assessment burger bun, the JuxtaLearn system supports the teachers in identifying a threshold concept and breaking it down into smaller, more manageable chunks or 'stumbling blocks'. The system then supports teachers as they enter Example Student Problems to describe the type of students' problems they've encountered as students try to understanding these threshold concepts. The Problem Distiller tool (see Figure 1) supports this process by prompting teachers to reflect on why these problems occur by making selections from a theoretically underpinned set of tabbed prompts. This detailed information about student problems and the possible causes of these problems is fed through to the quiz authoring tool to scaffold the quiz authors (teacher or student) as they write questions for the quiz.

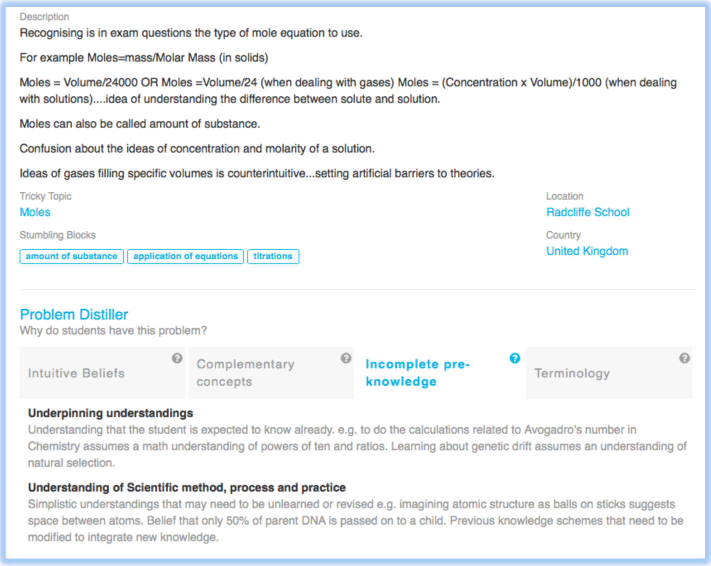
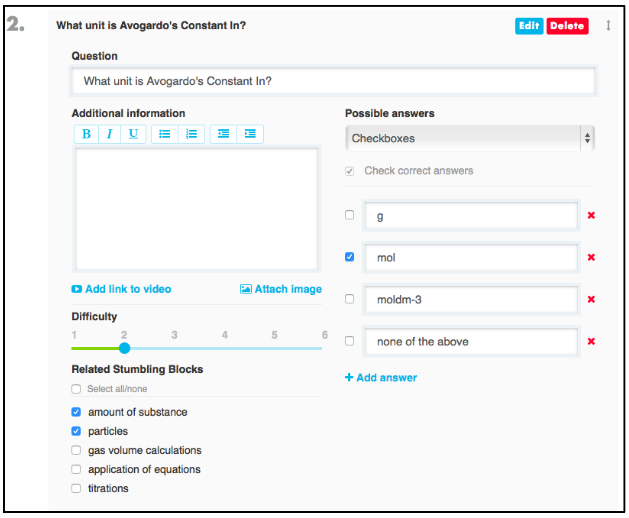



**Fig. 1.** JuxtaLearn Problem Distiller

Having identified the Tricky Topic and stumbling blocks, the teacher then moves on to write questions for the Diagnostic Quiz (see Figure 2). The first step when constructing a question in the JuxtaLearn quiz authoring system is to select the Tricky Topic and stumbling block(s) that the question is aimed at (see Figure 2a). This displays all the information related to that question; example student problems and suggested causes. From opening up section of the screen, all example student problems linked to stumbling blocks can be displayed along the bottom of the computer screen together with the Problem Distiller selections (see Figure 2b). This focuses the question authoring on probing the students' understanding rather than simply mapping directly onto a topic as taught in class.

(a)




(b)

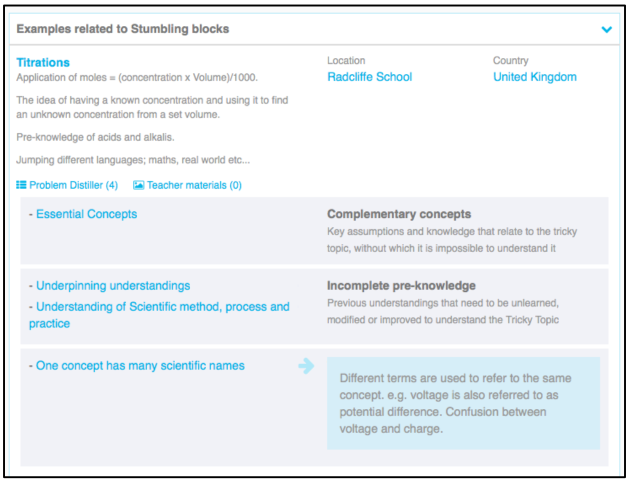



**Fig. 2.** Quiz Authoring and Problem Distiller Choices displayed during Quiz Authoring

*The bottom of the e-asssessment burger bun*

It is this connection to the stumbling blocks that determines the complexity weighting of each question, and which feeds into a radar chart visualisation for the student who has completed the quiz. The visualisation represents to the student their depth of understanding for that Tricky Topic. Complex questions may be linked to several stumbling blocks, whereas a question linked to only one stumbling block is generally simpler. The visualisations form the bottom of the e-assessment burger bun, enabling students to take ownership of their learning by targeting their knowledge gaps, and also informing teachers on areas of persistent student misunderstandings that could be addressed by a shift in teaching strategy.

Figure 3 illustrates how the JxL questions created by the teacher or students connect back to the Tricky Topic through the stumbling blocks they are tagged with. The Example Student Problems as entered into the system by the teacher, describing typical mistakes students make in class, are shown here connected to the stumbling blocks they illustrate. The Diagnostic Quiz questions are shown, weighted according to the stumbling blocks it is linked to.

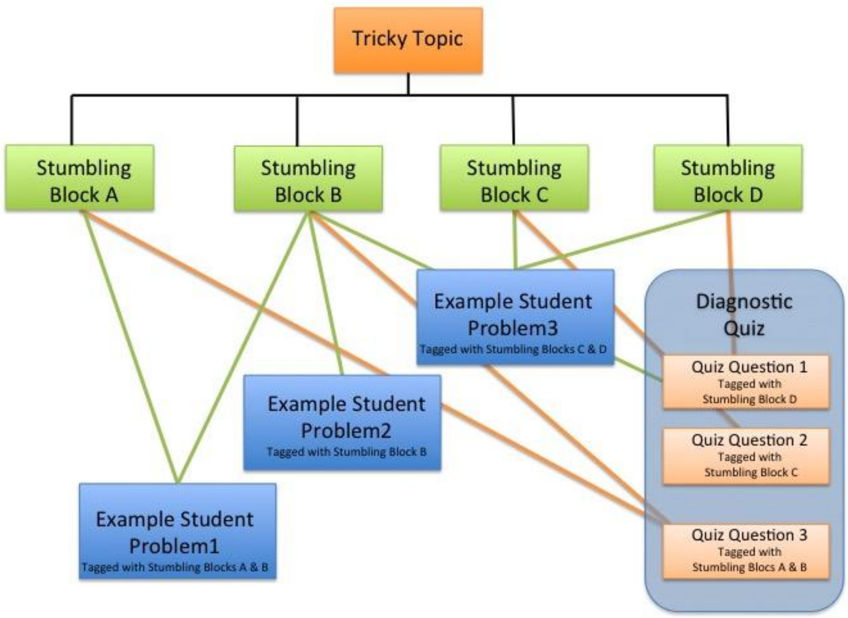



**Fig. 3.** JuxtaLearn Quiz Questions weighted using Stumbling Block tags

The next section on Methodology describes the trials conducted in schools with teachers and students and concludes with our findings to date.

## 2 Evaluation Methodology

A mixed methods approach was taken alongside a pragmatic epistemological approach to support a practice based underpinning to the evaluation activities. This approach was chosen to increase the practice based focus of the findings to increase the potential impact upon the schools eco-system.

The 3 trials described in this paper were run with UK teachers (from two different secondary schools) and their students. Table 2 details the descriptive statistics, as well as the data collection methods taken for these projects.

Table 2. UK Trials of the Diagnostic Quiz

| Workshop ID: UKRad1 | | Data Collection Methods | Data Analysis |
|---|---|---|---|
| Date: | 7/7/14 | • InterationVideo recordings<br>• Audio recordings<br>• Participant observation notes (livescribe pen recordings)<br>• Focus group teacher debrief video recordings<br>• Focus group student debrief (teachers)<br>• 2 Diagnostic Quiz results (Pre and Post quizzes)<br>• Student storyboards (paper)<br>• Student videos | • Thematic analysis of interviews and observation notes<br>• Interaction video analysis 3 staged approach:<br>1) critical incident identification,<br>2) thematic analysis of key incidents<br>3) interaction patterns mapped<br>• Pre and post activity quiz comparisons to identify ‘intervention’ impact<br>• Creativity and accuracy articulation mapping |
| Subject: | Drama | | |
| Participants | 11 students aged 16-18<br>1 teacher | | |
| Gender | 6M 5F | | |
| Workshop ID: UKRad2 | | | |
| Date: | 8/7/14 | | |
| Subject: | Chemistry | | |
| Participants | 8 students aged 16-18<br>1 teacher | | |
| Gender | 6F 2M | | |

| Workshop ID: UKSRA1 | | Data Collection Methods | Data Analysis |
|---|---|---|---|
| Date: | 20/5/15 | As above +<br>• Approaches to Learning questionnaire | As above +<br>• Approaches to Learning questionnaire analysis. Standardised questionnaire with standardised data groupings. |
| Subject: | Maths | | |
| Participants | 8 students aged 12-13 | | |
| Gender | 0M 8F | | |

### 2.1 Pre-Trial Preparation

For all three trials, the teacher logged onto the JuxtaLearn system in advance and prepared the materials for the student. These materials consisted of the Tricky Topic, stumbling blocks and Example Student Problems. Consent forms were obtained from all the students and their parents prior to the trial and information on the project was provided for the students. In both Trial UKRad 2, and Trial UKSRA1 the teachers initially worked alongside the researcher but once confident with the system, they finished creating the quiz in their own time. For UKRad2 , the Chemistry teacher adapted the process of quiz-writing, engaging his yr12 students to write the quiz questions for the yr11 students involved in the trial. With trial UKSRA1, the teacher created a slightly different quiz by using the same questions, but putting the multiple choice answers in a different order. In Trial UKRad1, the teacher logged on to the system alone without any support from the research team and prepared all the materials including the quiz unaided.

### 2.2 Trial Procedures

Trials 1 and 2 took place over the course of a full school day, Trial 3 was run as a series of 5 x 40minute sessions with the students during their school lunch break. All students took the quiz before forming into groups to undertake the storyboarding and video creation stages of the JuxtaLearn process. The students took the quiz again at the end of the day, or for Trial 3, during the last session. Whilst taking the quiz twice is not required by the JuxtaLearn process, it was requested by the teachers.

Before the JuxtaLearn learning activities began, the teacher and students reviewed their quiz visualisations and discussed the highlighted knowledge gaps. The teachers encouraged their students to focus their creative video making efforts on their biggest stumbling blocks. All students took the same quiz at the end of the trial to assess whether the intervention had helped them plug some of the gaps in their understanding.

In UKSRA1 trial, the students also took an online version of the Learning Process Questionnaire [30] before and after the intervention.

### 2.3 Analysis Methods

The analysis of the data was conducted using a thematic analysis approach combining systematic levels of abstraction into a model, which was verified through representations presented to the participants. In particular, the video and audio data were analysed using thematic analysis to identify; critical incidents, lightbulb moments and evidence of developing understanding. This data was supplemented with the data from the post-hoc interviews and focus groups. These then led to an expansion of the themes triangulated through the different data sources collected. A more specific meta-analysis was conducted on the assessment specific data sources. This was expanded upon by diagnostic quiz result and (for one trial) approaches to learning data from the students. Teachers feedback on the assessment creation and implementation procedure were then correlated with their perceptions of how these supported the whole teaching intervention (which was the whole Juxtalearn process). This process led to the development of a metaphor model with the ‘e-assessment’ burger representation. The before and after of assessment pedagogically connected to the learning process rather than simply to assess a learning outcome.

## 3 Findings and Discussion

The findings from the trials identified a wealth of data specifically focused upon the assessment tools and methods. As already detailed, the assessment related data were thematically analysed into three threads that were later reviewed as prior and post the e-assessment completion activity. This was later termed an e-assessment burger with the two ends connected to teaching and learning practice referred to as the 'bun' surrounding the e-assessment burger. It has been noted that frequently e-assessment systems focus on the technology with reference to the testing and marking aspects of the system. However, the results from these trials have highlighted the power of effectively relating formative assessment systems into teaching practices and personalised learning pathways.

As already noted, we review three issues in developing and evaluating the role of designing an eassessment system that is co-designed in connection to designing learning processes:

- Problems for / barriers to learning,
- depth of learning (deep and surface)
- reflection and learning ownership

In summary the findings identified that scaffolding formative assessment design can support not only teachers as authors for question creation but develop their understanding of students *barriers to understanding*. The system supported teachers reflection upon and accurate identification of gaps in students understanding. In particular, this was driven by the quizzes focus upon specific needs that the teachers were supported in identifing. This also gave teachers confidence in the system as an effective feedback mechanism. They then used the sysetm to guide follow on learning design, and teaching activities with students as co-creators of their learning in further question designs. This went beyond the original aims of the project and supported learning re-design based upon *the depth of learning* revealed by the assessment system. The e-assessment visualisations of results for whole cohorts was found to support teachers in identifying the effectiveness of their current teaching and potential ways to develop peer and student-directed learning, and their own input to future teaching activities. An unforseen effect on the agency of the system was the role of the system in supporting teachers professional development. Finally the most powerful aspect of the system design was around *the reflection and ownership* incured by the system. The systems intensive focus on question creation based on threshold concepts were found to increase students reflecting upon gaps in their learning increasing the value of assessment feedback for students. In particular this was found to shift ownership towards the student and thus a joint agency for teaching and learning naturally evolved with the e-assessment system as a lychpin in this process. However, it must be understood that this required a shift in approaches to assessment that once made changed conceptions of assessment. It could be argued that the eassessment burger is a threshold concept in itself that once understood could transform the role and agency of eassessment in the learning process.

### 3.1 Formative Assessment design: Scaffolding Question Creation

As was noted from prior research to support effective assessment that empowers student directed learning requires linking it to learning pathways. Within the Juxtalearn project a focus has been on scaffolding and supporting the learning process. Initially it is important with assessment tools to identify exactly what constitutes an effective assessment question. Within the evaluation trials this was found to depend, in part, on where the question occurred within the quiz. Teachers using the system noted that it was better to start out with simple questions that test the students' knowledge of the fundamental building blocks of the topic or threshold concept. The teachers when authoring the quiz were also found to review question sequencing that uncovered surface and deep learning. In particular, within this quiz design the initial questions devised were relatively simple and could often be answered by students using surface learning with memorised information. As the quiz progressed, students are confronted with exploratory questions that dig deeper and asked the student to apply the knowledge that they had memorised. Towards the end of the quiz the students were presented with more complex questions, often with multiple stumbling blocks associated, that drilled down to reveal whether or not students had that deeper understanding of the topic or the ability to apply their knowledge in different contexts.

The layers of understanding was supported in the authoring tool through the author identifying and assigning one or many stumbling blocks when writing a quiz question. This provides the learner with a variation in question, expressing different depths of understanding from surface to deep. However, the structured sequencing of simple to difficult questioning was not directly supported in the quiz authoring tool but has been considered for later versions.

Although the project had not set out to support students creating quiz questions the pedagogical scaffolding for quiz creation was opportunistically found to support this activity. Within trial UKRad2 the teacher had identified the Tricky Topics, stumbling blocks and student problems they felt were required for an understanding of the topic. One part of the process they found particularly helpful was the step of identifying 'why' students were encountering problems in understanding these concepts. These accounts, coupled with the systems scaffolding with problem distiller prompts deconstructing 'why' this maybe a problem for the students increased support for the quiz question author. This was found to be so facilitative that one teacher felt comfortable enough to direct his older students to create questions for the year below them. This was not directed by the research team.

Although the teacher was semi-driven by pragmatics of poor time resources, he also noted the value of students testing their understanding through question creation. It is interesting to note that the teacher initially wrote the first set of pre-quiz questions (see table 2) himself and once he was comfortable and saw the value of the system, the structured support and procedures he decided to give it as a task for his 2$^{nd}$ year 6$^{th}$ form students for the 1$^{st}$ year students to take. Within the feedback procedure he noted that the students who constructed these quizzes had enjoyed the process and interestingly enough the students who took the second quiz (created by the students from the year above) said it was harder than the first quiz (created by their teacher). Having said that the interim JuxtaLearn focused learning had increased their deeper understanding of the concepts. For learning evaluation purposes the scaffolding of the quiz creation also worked well. This is because the threshold concepts, stumbling blocks and weighting remained the same for both the quizzes allowing for more effective direct comparison between the two quizzes.

### 3.2 Taking the Quiz: E-assessment supports student ownership of learning

The evaluation identified that most students who completed the simple questions with a surface knowledge of the threshold concept struggled with questions requiring a deeper understanding. Within the trial UKRad2 students responded that they did not like the Diagnostic Quiz because it had harder questions: *"When we did questions, when we were doing it out of the book, [...] and I was completely understanding some of the questions [...] but I did a question on the quiz and I had to, like do it, apply it differently, and then I got it wrong". (Chem4)*

In trial UKSRA1, when providing the students with feedback on their results they noted that this gave them with more clarity on their depth of understanding, as one maths student noted: "*Some questions were very easy but some were very hard and I had to guess them*." *(SRAMath3)*

The observational data identified that students who took the diagnostic identified gaps in their understanding and then focused their further learning activities within the Juxtalearn process on these gaps. One chemistry teacher pointed out: *"Having the quiz for sure at the beginning focuses the [learning activities] filming, the video making." (ChemistryTeacher1)*

The formative assessment feedback mechanism gave students a standard percentage success score. This presented an indication of how well they had done, but did not highlight specific areas of misunderstanding. For example, Table 3 shows the before and after Diagnostic Quiz results for a class of chemistry students. These students took the same quiz before and after the JuxtaLearn intervention. Although the percentage scores show whether or not the student marks improved, and if so by how much, it is not immediately apparent where these changes in understanding occurred.

Table 3 Chemistry Diagnostic Quiz Results as a Percentage

| Chemistry Group | Clipit ID | Pre Quiz | | Post Quiz | |
|---|---|---|---|---|---|
| Group1 | Chem1 | 5/10 | 50% | 6/10 | 60% |
| | Chem2 | 4/10 | 40% | 7/10 | 70% |
| | Chem6 | 5/10 | 50% | 5/10 | 50% |
| | Chem7 | 7/10 | 70% | 8/10 | 80% |
| Group2 | Chem5 | 3/10 | 30% | 7/10 | 70% |
| | Chem3 | 5/10 | 50% | 9/10 | 90% |
| | Chem8 | 4/10 | 40% | 5/10 | 50% |
| | Chem4 | 2/10 | 20% | 3/10 | 50% |

Even presented with a breakdown of which questions they had got right and which ones were wrong it can still be difficult to pinpoint the precise gaps in understanding. For this, the JuxtaLearn quiz radar chart visualisation is key. This visualisation presents students' results in terms of how they performed in their understanding of the threshold concept stumbling blocks (see Figure 2). The radar chart visualisation makes it clear where the gaps in knowledge are, and where any changes in understanding has occurred as the student works through the JuxtaLearn learning activities (see Figure 4). Because the questions are linked to the stumbling blocks, with simple questions linked to only one stumbling block and more complex questions linked to several stumbling blocks (see Figure 3), the visualisation gives a good indication of the levels of understanding and is easier to interpret for both students and teachers. Figure 4 shows the before and after visualisations for chemistry student Chem3. The before quiz visualisation shows a surface (poor) understanding of **gas volume calculations** and **application of equations**. The 'after' the juxtalearn intervention visualisation shows that the student improved after focusing their learning on these weaker areas, developing a deeper understanding of the concept.

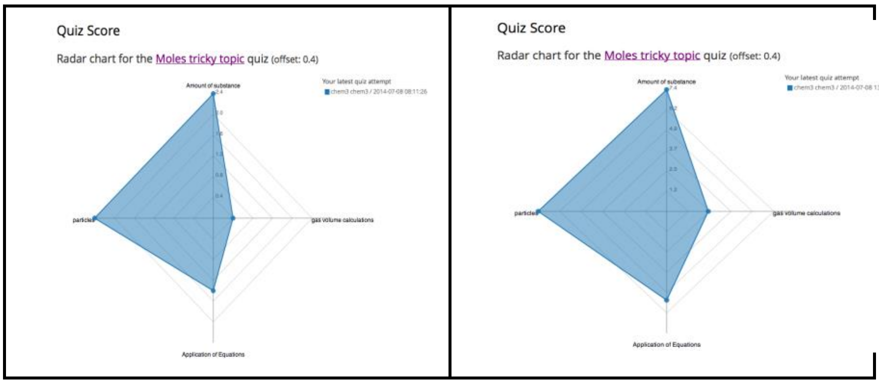



**Fig. 4.** Before & After Radar Chart Quiz Visualisation UKRad2 (Chem3)[2]

Within trial UKRad2, students used the visualisation to direct where they should focus their future Juxtalearn learning activities to develop a deeper understanding of the concept. Whilst students were found to be quite competitive with each other when completing the questions, the visualisation produced a more personally reflective approach from the student on their levels of understanding. This role of the Diagnostic Quiz visualisation facilitated a shift in agency from classroom competition to student reflection and learning ownership which occurred across different subjects. Several of the teachers across the trials identified during post-workshop debriefing sessions that they noticed this shift in learning approach. For example, after the UKRad1 trial in which the teacher had identified the topic of theatre genres as the problem concept for his students, he commented: *"One of the students seemed to make no progress in terms of the quiz [*in terms of her marks*] , for that she felt disheartened. I showed her, her visualisation and she could see that one of the areas that she hadn't done well, a big gap, was what she felt out of the entire thing that she was strongest on, so she has gone home to bone up on Epic Theatre. She's realised that she wasn't as secure as she felt she was, and the quiz has highlighted that. That's priceless, absolutely priceless." (DramaTeacher1)*.

[2] These visualisations show broad visual overviews of shifts in learning with related stumbling block titles (e.g. particles) not as valuable for these purposes so not clearly depicted in the image.

Across all the trials the findings identified that it was the tight pedagogical connection between the quiz questions and the teacher-identified threshold concepts that transformed this assessment into a tool for change. This approach enabled students to take ownership of their learning with more detailed information on the level of their ‘deep’ understanding and how this was progressing. The quiz visualisations also enable students to personalise their learning pathway, identifying and adopting strategies that focused on the gaps in their knowledge.

### 3.3 Quiz feedback: visualisations supporting teaching practices

The quiz visualisation feedback was initially envisaged as a mechanism to support students interpreting their depth of understanding for the threshold concept and related stumbling blocks. The tool also offered visualisation of the cohort of results overlayed on each other for the teacher. The evaluation identified that these cohort visualisations supported a quick overview of that groups’ understanding for the teacher (see Figure 5). This was popular with the teachers. *“I think it was useful for a snapshot view of the group”. (ChemistryTeacher1)*

However, the evaluation identified that the teachers’ usage of the data from the quizzes became far more nuanced because of the class-wide visualisation. For example one teacher noted impacts on their future teaching practices, as well as the students own directed learning activities based upon the quiz visualisation.

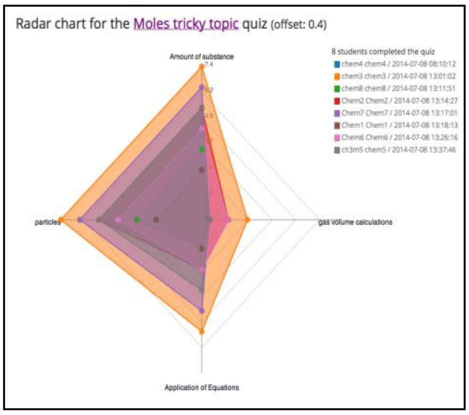



**Fig. 5.** Chemistry Post-Quiz visualisation[3]

The teacher deconstructed that the visualised results for particles in Figure 5 identifying ways to pair learners for peer learning activites. “*It makes me go maybe that student could teach that student and try and unpack their learning and then help these ones out as well. This makes me think the intervention could be with them.” (ChemistryTeacher1)*

[3] These visualisations show broad visual overviews of shifts in learning with related stumbling block titles (e.g. particles) not as valuable for these purposes so not clearly depicted in the image.

In addition, the same teacher identified that in contrast for gas volume calculations there was a more generic issue of understanding across all the students identified: *"But this [gas volume calculations] has to be a bit more me. I've then got to go look at it again and go, what is it about that ... it could be me... possibly re-teaching it again could be useful." (ChemistryTeacher1)*

The whole quiz interpretation process produced a reflective approach for the teacher. This was initially because the visualisation was tied to the learning the teacher had defined through Tridky Topics and stumbling blocks. This information on Tricky Topics and stumbling blocks, when combined with the resulting visualisations increased the potential for meaning making from the visualisations. As a result of the pedagocial underpinning the first steps in the quiz making, the visualisations became a powerful tool in re-designing subsequent learning activities according to the students' personal needs.

The e-assessment system feedback was also found to inspire teachers reflections on the sequencing of their teaching, where formative assessment fitted within the learning process and how to improve subsequent versions of this e-assessment system: *"Making the quiz better, in terms of having the post-quiz being specific to the area they drop in their knowledge, the stumbling block, the very specific stumbling block." (ChemistryTeacher1)*

Ultimately the evaluation findings identified that the visualisations triggered a more effective process of reflection for the teacher, supporting understanding individual student's needs as well as the whole cohort's understanding. Feedback on whole classes were noted as pinpointing areas of persistent student misconception that triggered changes to teaching activities and ultimately teaching practices.

It is interesting that one finding from the project is that across all the schools where this project was conducted, the impacts are felt across the whole school. In one school this caused an element of confrontation across the hierarchy in the teaching staff. In particular this was identified as due to the fundamental changes to concepts to e-assessment and teaching practices that it triggered. Whilst for many this is a welcomed catalyst for change to the educational eco-system, for others it is yet another upheaval to established social norms and practices.

## 4 Conclusions

The formative assessment findings from this project have been broken down into three main themes all of which relate to empowering agency changes at a 12-18 yr old school level. Initially scaffolding the question design system around teacher defined threshold concepts and related stumbling blocks provided an increased confidence in the value and relevance of the formative assessment system. This approach was found to increase the effectiveness of developing specific questions, support some authors in reflecting upon and developing a sequenced approach to the complexity of question levels. The threshold concept stumbling block approach also provided a way to scope depth of understanding in questions. This combined with the Problem Distiller supported effective question creation to the level that teachers felt not only comfortable, but the value of students creating questions.

When students took the formative assessment quizzes they themselves were found to value not the simplistic percentage feedback but the radar visualisations. It was not the visualisation on its own that was of value to them but its ability to reflect their depth in understanding at a deep or surface level. This was achieved through weighting assigned via stumbling blocks and the number of these assigned at the question creation stage. The teachers also found this level of feedback valuable as they used it to understand not only in more depth individual gaps in understanding but as a cohort where the students were requiring additional teaching and learning.

Supporting teachers to identify threshold concepts ('Tricky Topics' as they have termed them) and related stumbling blocks and problem examples provides a focus for question writing whether that be by the teachers or the students. As was noted by one teacher the system allowed him confidence to assign question writing to his students for the year below to complete. This was highlighted as valuable not only to support those taking the quiz, but those creating the questions. The level of import for this teacher was highlighted by how valuable and limited teaching time is within a secondary level science subject. Yet this teacher still found student created quiz questions as valuable enough to incorporate.

The assessment tools and procedures developed within this project have been developed as a means to empower learners to take up student-directed learning approaches with the Juxtalearn Process. The Diagnostic Quiz was found to support a shift in agency, embedding assessment as part of a student-directed learning process either by directing flipped learning or through earlier reflections on quiz question creation for fellow students. Through mapping the assessment systems to a wider curriculum and learning pathway purpose the project identified a broader and more powerful role for assessment. Teachers became creative and enthused by formative assessment systems, empowered to use and re-invent how it was implemented to support their teaching practices.

This formative assessment with associated visualisations also gave students ownership of their learning, allowing them to focus on gaps and see assessment as a tool to support and reinforce their learning. Student started to move beyond assessment as an end goal, into identifying with it as a tool to help them direct their own learning activities. Evaluation activities have captured video accounts of teachers shocked by their own excitement in using and developing assessment systems. "*I can't believe I'm enjoying using assessment"* (*DramaTeacher1*).

Another teacher closely involved with this project became so enthused by assessment that he became an advocate for it within his school and was promoted to be head of assessment (a new role within the school) across subjects. One central theme that has driven this teacher in all of his activities has been the effectiveness of feedback mechanisms, not only for the students in interpreting this feedback but for the teachers in effectively providing this feedback and adapting their teaching practices according to student needs. However, within one school this triggered a minor clash between teachers around concepts of the role of assessment and teaching approaches. It should be noted that when trialling and evaluating assessment innovations that changing advocacy with the formal educational ecosystem can be both threatening for some as well as empowering for others.

As the title for this paper denotes this project evaluation presents an e-assessment burger. The initial formative assessment creation tied into curriculum teaching is the top half of the bun, taking the quiz are the inside burger, whilst the results feeding into identifying further learning is the bottom half of the burger. All too often we have become focused on the assessment core, taking and scoring the results of the quiz. However, within the tie into learning design in creating the quiz and back to further learning activities upon receiving the results we let loose the real value of formative assessment systems. It is through these connection points that assessment systems can change the agency, relationship and role of teachers and students. All too often assessment has become a yardstick to jump over or simply to beat someone down into becoming a demoralised learner. This research has identified how connecting the before and after of formative assessment tools can change its role within education and effect a change in the whole ecosystem of education. Although, as has been highlighted, this requires a shift in perceptions of the role of eassesment in the learning process. The eassesment burger supports changing concepts of the role of eassessment. However, it could be argued that the eassessment burger is a threshold concept in itself that is difficult to comprehend but once understood could transform the role and agency of eassessment in the learning process.

## References


1. Crisp, G.: The e-Assessment Handbook. London, Continuum, (2007)
2. Jordan, Sally.: E-assessment: past, present and future. New Directions, 9 (1), pp. 87–106, (2013)
3. Whitelock, D. M. and Brasher, A.: Developing a Roadmap for e-Assessment: Which Way Now? In: Danson, Myles ed. Proceedings of the 10th International CAA Conference. Loughborough, UK: Professional Development, Loughborough University, pp. 487–501, (2006)
4. McAllister, D. & Guidice, R.M.: (2012) This is only a test: A machine-graded improvement to the multiple-choice and true-false examination. Teaching in Higher Education 17 (2), pp. 193–207, (2012)
5. Waights V.: Assessing learning for healthcare practice using an on-line interactive tool. In: NET Conference, 8-10 Sep 2009, Cambridge, UK, (2009)
6. Bacigalupo, D. A., Warburton, W. I., Draffan, E. A., Zhang, P., Gilbert, L., & Wills, G. B. A formative eAssessment co-design case study. In Advanced Learning Technologies (ICALT), IEEE 10th International Conference pp. 35-37, IEEE.,(2010).
7. Fulton, K.: Upside down and inside out: Flip your classroom to improve student learning. Learning & Leading with Technology, 39 (8), pp. 12–17, (2012)
8. Herreid, C. F. & Schiller, N. A.: Case studies and the flipped classroom, Journal of College Science Teaching 42 (5), pp. 62-66, (2013)
9. Ashton, H.S. and Thomas, R.C.: Bridging the gap between assessment, learning and teaching. In Proceedings of the 10th International CAA Conference, Loughborough, (2006)
10. Gipps, C.V.: What is the role for ICT-based assessment in universities? Studies in Higher Education 30 (2), pp 171–180, (2005)
11. Entwistle , N.: Styles of Learning and Teaching, David Fulton, (1988)
12. Biggs, J. B.: Teaching for Quality Learning at University, SHRE and Open University Press. 1999

13. Laurillard, D.: Rethinking University Teaching, a framework for the effective use of educational technology, Routledge, (1993)
14. Mezirow, J.: A critical theory of adult learning and education, Adult education quarterly 32 (1), pp. 3-24, (1981)
15. Meyer, J. & Land, R.: Threshold concepts and troublesome knowledge: linkages to ways of thinking and practising within the disciplines. In Rust, C. (Ed.) Improving student Learning - Theory and Practice Ten Years on. Oxford, Oxford Centre for Staff and Learning Development (OCSLD), pp.412-424, (2003)
16. Meyer, J. & Land, R. Overcoming barriers to student understanding: Threshold concepts and Troublesome Knowledge. In Meyer, J. & Land, R. (Eds.) *Overcoming Barriers to Student Understanding: Threshold concepts and Toublesome Knowledge.* London and New York, Routledge, pp.19-32, (2006)
17. Harlow, A., Scott, J., Peter, M. & Cowie, B.: Getting stuck" in Analogue Electronics: Threshold Concepts as an Explanatory Model. European Journal of Engineering Education, 36 (5), pp. 435-447, (2011)
18. Lucas, U. & Mladenovic, R.: The potential of threshold concepts: an emerging framework for educational research and practice. *London Review of Education,* 5 (3)**,** pp. 237-248, (2007)
19. Machiocha, A.: Teaching research methods: threshold concept. In 13th European conference on research methods for business and management, London, (2014)
20. Rowbottom, D. (2007) Demystifying Threshold Concepts. *Journal of Philosophy of Education,* **41** (2)**,** 236-270.
21. O'Donnell, R. A Critique of the Threshold Concept Hypothesis and an Application in Economics. *Working Paper Series, 164,* (2010) Available at http://www.finance.uts.edu.au/research/wpapers/wp164.pdf [Accessed 31-3-14],
22. Denny, P., Luxton-Reilly, A., Hamer, J., & Purchase, H.: Coverage of course topics in a student generated MCQ repository, In *ACM SIGCSE Bulletin* 41 (3), 11-15, (2009)
23. Bull, J. & McKenna, C.: Blueprint for computer-aided assessment. London: Routledge Falmer, (2004)
24. McGuire, G.R., Youngson, M.A., Korabinski, A.A. & McMillan, D.: Partial credit in mathematics exams: A comparison of traditional and CAA exams. In Proceedings of the 6th International Computer-Assisted Assessment Conference, University of Loughborough, (2002)
25. Ihantola, P., et al.: Review of recent systems for automatic assessment of programming assignments. Proceedings of the 10th Koli Calling International Conference on Computing Education Research. Koli, Finland, ACM**:** 86-93, (2010)
26. Bacon, D.R.: Assessing learning outcomes: A comparison of multiple-choice and short answer questions in a marketing context. Journal of Marketing Education 25 (1), pp. 31–36, (2003)
27. IMS Global Learning Consortium IMS question and test interoperability specification, (2013) Available at http://www.imsglobal.org/question/ (accessed 10 May 2015).
28. Schulte C. & Bennedsen J. (2006). What do teachers teach in introductory programming? In ICER '06: Proceedings of the 2006 International Workshop on Computing Education Research, pages 17–28, New York, NY, USA, 2006.
29. Luxton-Reilly, A., et al.: Supporting student-generated free-response questions. Proceedings of the 16th annual joint conference on Innovation and technology in computer science education. Darmstadt, Germany, ACM: 153-157, (2011)
30. Biggs, J. B.: Student Approaches to Learning and Studying. Research Monograph, Hawthorn, Australia, Australian Council for Educational Research Ltd, (1987)