

CASE STUDY

Embeddedness of mathematics e-assessment and attitudes affecting adoption

Deirdre Casey, Department of Computing, Griffith College Cork, Cork, Ireland.

Email: Deirdre.Casey@griffith.ie

Julie Crowley, Mathematics Department, Cork Institute of Technology, Cork, Ireland.

Email: Julie.Crowley@cit.ie

Abstract

The Numbas e-assessment system was first introduced to Cork Institute of Technology (CIT) in 2014. It has been implemented in several modules and is used regularly by a number of Lecturers. The aim of this case study was to investigate how embedded Numbas is in the Mathematics department in CIT and to examine the perceptions of the lecturing staff which have led to their individual adoptions of the Numbas e-assessment tool. The results of a survey of lecturers within the department show that the use of Numbas is common and varied and that Lecturers' attitudes towards Numbas are positive. Several themes arose around ways to increase or improve the use of Numbas including the need for ongoing training and ideas for the broadening of use. Lessons learned could be applied to advancing the integration of Numbas in Cork Institute of Technology but may also be of use to others hoping to increase adoption of e-assessment in Mathematics within their institutions.

Keywords: Mathematics, e-Assessment, diffusion of innovations, implementation, computer aided assessment

1. Introduction

In 2014 Cork Institute of Technology (CIT) introduced e-assessment into some of its Mathematics modules. The initial motivation for using e-assessment was to manage lecturer workload in correcting continuous assessments thereby allowing more regular assessment and feedback. The Numbas e-assessment tool was chosen because it has a strong reputation, is user friendly and is compatible with the Virtual Learning Environment used on campus. The National Forum for the Enhancement of Teaching and Learning funded a joint project between University College Cork and Cork Institute of Technology in 2015 called 'Transitioning to e-Assessment in Mathematics Education (TEAME)'. The project focused on the implementation and evaluation of this Mathematics e-assessment tool and the dissemination of associated knowledge. Along with the expected benefits of facilitating more regular assessment and immediate feedback Numbas also increased student engagement and attendance in their tutorials at CIT (Carroll, et al., 2017). Numbas was rolled out to several large Mathematics modules and became integrated in the module delivery. The large scale and departmentally supported introduction of Numbas in CIT and its relative success could be used to help other institutions to encourage greater levels of adoption within their ranks.

The aim of this research is to evaluate the embeddedness of Numbas e-assessment in the department and to investigate the lecturer attitudes which have led to this level of adoption. This research focuses on the lecturer perspective rather than that of the student and uses lecturer opinion to evaluate levels of embeddedness. Qualitative and quantitative data was collected through a survey in order to ascertain lecturer perceptions of the scale and width of Numbas use and their attitudes towards adopting Numbas.

The remainder of this paper is organised as follows. The next section discusses the literature in relation to diffusion of innovations and ways to evaluate the embeddedness or success of

implementation. Next the methodology and methods used are considered. Then the results of the case study are presented. The final section discusses these results and their implications for Cork Institute of Technology and other institutions hoping to successfully implement Mathematics e-assessment.

2. Literature Review

The literature review focused on the spread of new ideas and how this knowledge can be applied to understanding the spread of Mathematics e-assessment.

2.1 Adoption and Diffusion of Innovations

Every new idea starts with an individual or small group and gradually diffuses outwards from that small starting point. *Adoption* is the process whereby an individual goes from first hearing about an innovation to finally adopting it. According to Rogers (1962) the adoption process can be viewed as five consecutive stages: awareness, interest, evaluation, trial and adoption. The spread of the adoption of new ideas is described by Rogers (*ibid.*) as *Diffusion of Innovation*. He described five user perceptions that effect the diffusion of any new innovation: relative advantage, compatibility, complexity, divisibility, communicability. These attitudes from Rogers work were refined and formed into an instrument to measure perceptions of adoption by (Moore & Benbasat, 1991) which measures eight perceptions:

- *Voluntariness* is the degree to which use of the innovation is perceived as being voluntary, or of free will;
- *Relative Advantage* is the degree to which an innovation is perceived as being better than its precursor;
- *Compatibility* is the degree to which an innovation is perceived as being consistent with the existing values, needs, and past experience of potential adopters;
- *Image* is the degree to which use of an innovation is perceived to enhance one's image or status in one's social system;
- *Ease of Use* is the degree to which the innovation is perceived as being easy to use;
- *Result Demonstrability* is the degree to which the results of using an innovation are perceived to be measurable, observable and easy to communicate to others;
- *Visibility* is the degree to which the innovation is perceived to be visible within the organisation/department;
- *Trialability* is the degree to which it is perceived that the innovation can be experimented with before adoption.

High scores in these eight perceptions predict adoption of the innovation by an individual.

2.2 Mathematics e-Assessment and its Diffusion

e-Assessment is assessment that uses or is enhanced by electronic technologies and it can offer many advantages over traditional pen and paper assessment. These advantages include immediate feedback, the possibility for students to make more than one attempt and supplying hints (Biggs & Tang, 2007; Jenkins, 2004). *e-Assessment* in Mathematics gained momentum in the 1990's with the development of WeBWork at the University of Rochester and has continued to grow and diversify. In 2012 developers at Newcastle University released a new open source Mathematics e-assessment system called 'Numbas'. Numbas is now used successfully in several large institutions around the world (Loots, et al., n.d.; Foster, et al., 2012).

Although e-assessment in Mathematics is far from a new innovation its diffusion rates through the mathematics education community remain slow. The uptake of Computer Aided Assessment (CAA) has lagged behind the expectations of academics in the field (Warburton, 2009). This may be due to the fact that apart from the advantages that e-assessment may offer it also comes with difficulties and potential risks (Jenkins, 2004). A major barrier to the implementation of CAA is the time investment required in learning how to use the system and also in the development of questions (Jenkins, 2004). Staff require training on how to use the particular system along with general guidance on good question design (Sim, et al., 2004). The level of initial and ongoing staff training required relies, at least in part, on institutional support which seems to be very beneficial in terms of ensuring implementation of CAA (ibid.).

Judging or measuring the level of embeddedness of e-assessment has been considered by several authors. Metrics for successful implementation were discussed by Warburton (2009) but he doesn't describe an instrument to measure these metrics. However, McCann (2010) uses a questionnaire to attempt to measure these metrics and we have followed a similar approach. Here we measure 'Width' of practice by focusing on range of use and number of users (Warburton, 2009). A question was included on the survey to investigate the range of uses of the Numbas system in terms of the three types of assessment: Assessment of Learning (AoL), Assessment for Learning (AfL) and Assessment as Learning (AaL) (Zeng, et al., 2018).

3. Methodology

This case study was carried out through mixed methods research using a concurrent embedded design (Creswell, 2009). Most of the data collected was quantitative with embedded qualitative elements where the qualitative data is intended to provide a supporting role to the quantitative data. The two methods are used side by side to provide different perspectives on the same case. The survey used contained four parts: Demographics, width of use and embeddedness, perceptions of adoption, open-ended questions. The third section of the survey, dealing with perceptions of adoption, was adapted from the 25-item instrument developed and recommended by Moore and Benbasat (1991). The respondents were requested to rate their agreement to each statement on a seven-point scale, 7 = extremely agree and 1 = extremely disagree. The survey ended with two open ended questions for the purposes of collecting supporting qualitative data. The questions were as follows: (i) What do you think would increase or improve the use of Numbas within CIT? (ii) Do you have any other comments relating to the usefulness or adoption of Numbas?

4. Methods

The survey was sent to the current members of the Mathematics Department in CIT. Quantitative data analysis was carried out on the data from the first three sections of the survey. The mean and standard deviation were found separately for each item in the Perceptions of Adoption section. The data was then grouped by perception and an overall mean and standard deviation was found for each perception.

The responses to the open-ended questions were combined due to the similarity and overlap of responses. The text data was subjected to conventional content analysis where codes are derived directly from the data without reference to pre-conceived categories. Directive content analysis was also undertaken to attempt to code the data in relation to the eight perceptions described above but the data did not fit well into these categories and so the researchers felt that this would not give a valid representation of the data obtained.

5. Results

The survey received 9 responses from the 29 lecturers in the department giving a response rate of 31%. All respondents had some experience working with Numbas and no responses were received from non-users of the system.

5.1 Width of Practice/ Metrics for successful implementation

Figure 1 shows that respondents have used Numbas for all three types of assessment:

- Assessment of Learning (AoL): Testing prerequisite knowledge, Diagnostic testing;
- Assessment for Learning (AfL): Low stakes assessment and homework;
- Assessment as Learning (AaL): No stakes assessment in tutorials, Support tool.

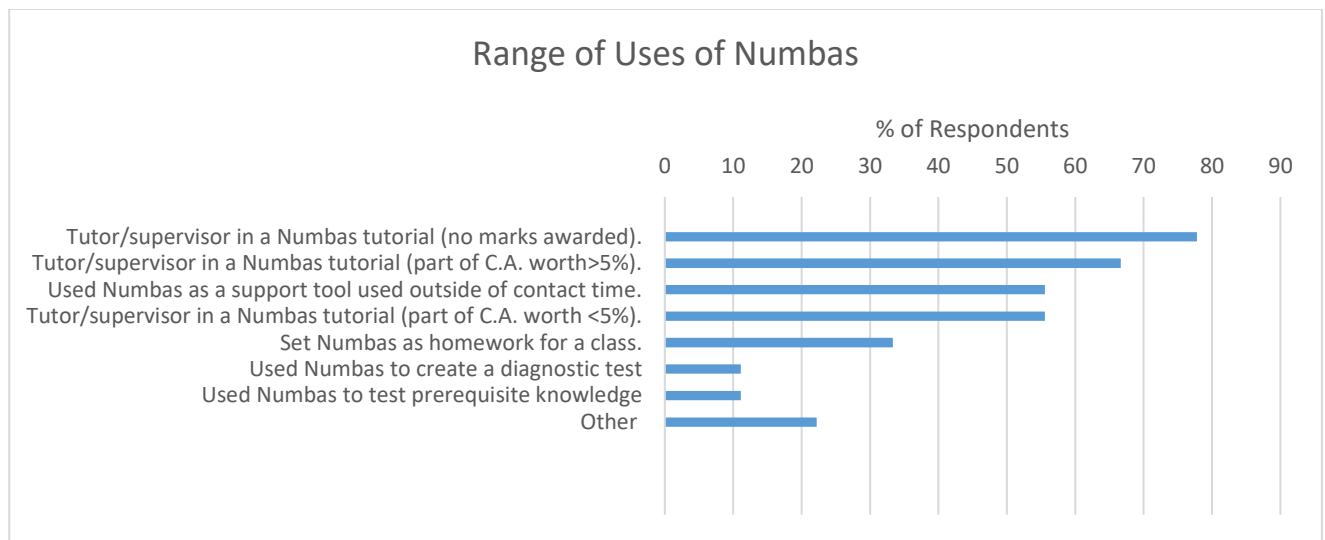


Figure 1: Range of Uses of Numbas among Respondents

5.2 Depth of use/knowledge

As shown in Figure 2 the respondents reported varied levels of engagement with the Numbas system. Over 50% of respondents have developed Numbas questions and so they are actively contributing to the question bank. Only one respondent had not done any of the listed activities using Numbas.

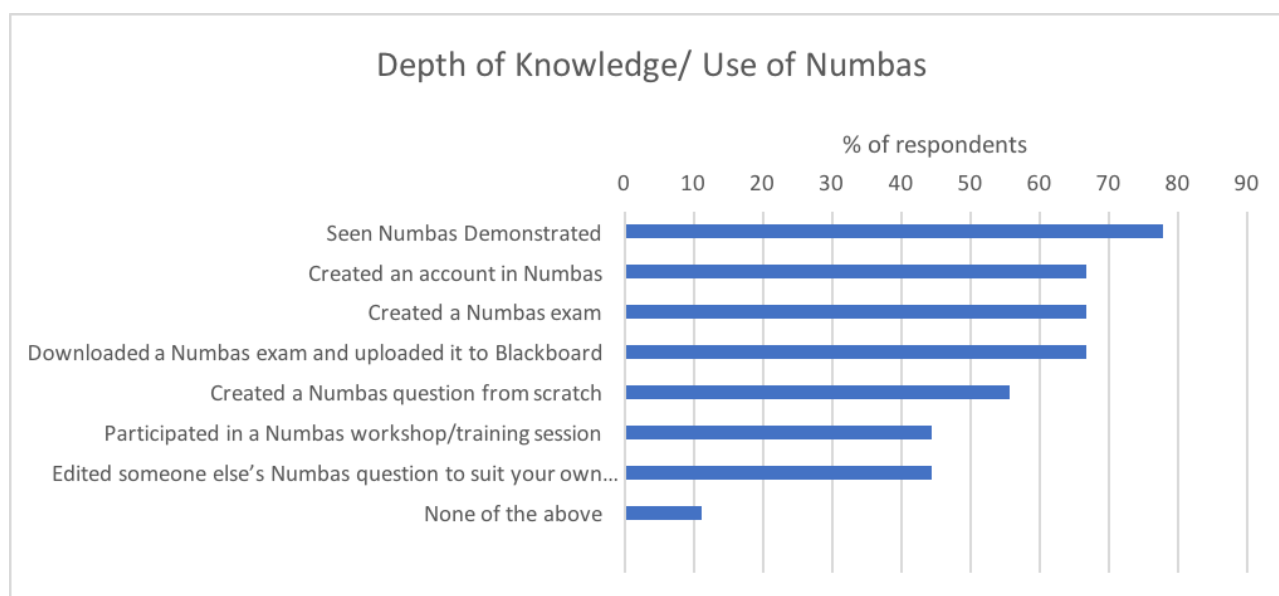


Figure 2: Depth of Use of Numbas Among Respondents

5.3 Attitudes/Adoption Attributes (of adopters)

The results of the Instrument to Measure the Perceptions of Adopting Numbas e-assessment are summarised in Table 1 where the respondents had rated their agreement to each statement on a seven-point scale, 7 = extremely agree and 1 = extremely disagree. The most positive perception was “Results Demonstrability” while the least positive was “Image”. Trialability and Image were the only two perceptions that did not elicit agreement from the respondents.

Table 1: Overview of the Perceptions of Adopting Numbas: Where respondents were asked to rate their agreement with statements related to each attribute on a seven - point scale 7 = extremely agree and 1 = extremely disagree.

Attribute	Mean	SD	Scale of Agreement
Results Demonstrability	5.81	1.21	Agree
Voluntariness	5.5	2.26	Agree
Compatibility	4.92	1.66	Somewhat Agree
Ease of Use	4.85	0.88	Somewhat Agree
Visibility	4.69	1.31	Somewhat Agree
Relative advantage	4.64	1.21	Somewhat Agree
Trialability	4.37	2.17	Neutral
Image	2.11	1.15	Disagree

5.4 Qualitative Data

All respondents gave a response to at least one of the two open ended questions: (i) What do you think would increase or improve the use of Numbas within CIT? (ii) Do you have any other comments relating to the usefulness or adoption of Numbas? Five concepts presented themselves from thematic analysis:

- The need for training;
- Expansion of use;
- Fairness and transparency to students;
- Need for resources;
- Teaching and Learning concerns or considerations.

Of the five concepts the need for training was the most prevalent. The desire for more training was alluded to by 5 of the 9 respondents. Within the five key concepts there were thirteen categories. These are summarised in Figure 3.

Concepts and Categories

Training	Expansion of Use	Fairness and Transparency	Resources	Teaching and Learning
<ul style="list-style-type: none">• Workshops• Ongoing training• Learning from each other	<ul style="list-style-type: none">• Low stakes homework• More assessments• More modules• Spread the word	<ul style="list-style-type: none">• Harsh on students• Data loss	<ul style="list-style-type: none">• Effort and time• Question bank• Funding	<ul style="list-style-type: none">• Evidence base• Other Tools

Figure 3: Five Main Concepts and thirteen categories emerging from the qualitative data

Table 2 provides a detailed breakdown of the concepts and categories along with illustrative quotes from the data for each category.

Table 2: Concepts, Captions and Supporting Quotes

Concepts	% of 23 items	Categories	Sample Items
Training	26.09% (6 items)	Workshops	<i>"Training on setting up and writing questions in Numbas"</i>
		Ongoing Training	<i>"a workshop focusing on the creation of more advanced questions"</i>
		Learning from each other	<i>"an experienced user and a new user [could] develop content for a course that they are both teaching together"</i>
Expansion of use	29.09% (6 items)	Low Stakes Homework	<i>"could potentially be used for homework too"</i>
		More Assessment	<i>"more assessments via Numbas"</i>
		More Modules	<i>"more accessible to other Mathematics Modules rather than basic first year modules"</i>
		Spread the Word	<i>"introducing it to a larger number of staff"</i>
Resources	21.74% (5 items)	Time and Effort	<i>"Time to develop more questions – writing new questions is time consuming"</i>
		Question Bank	<i>"Bigger question banks"</i>
		Funding	<i>"Funding"</i>
Fairness and transparency	13.04% (3 items)	Harsh on Students	<i>"It doesn't take into account for people that get correct answer but could be off by a decimal place"</i>
		Data Loss	<i>"being able to see the students' responses to the questions"</i>
Teaching and learning	8.7% (2 items)	Evidence Base	<i>"Evidence of students using Numbas having improved learning"</i>
		Other Tools	<i>"I think it is a good tool, but just one of many different tools you can use to teach"</i>
Other	4.35% (1 item)		<i>"I haven't used Numbas yet: still developing tools"</i>

6. Discussion

The results show a broad width of practice using Numbas within the department both in terms of range of use and number of users. Together these results indicate that Numbas has become reasonably well embedded in the department. All the respondents have had some level of engagement with Numbas. They all fall somewhere between the evaluation, trial and adoption stages of Rogers stages of adoption of an innovation (Rogers, 1962) tending mostly towards the adoption phase. No conclusions can be reached from our results for perceptions or attitudes of non-adopters or reasons for not adopting. Among responders the range of uses of Numbas were large and varied. It is used for Assessment of Learning (AoL), Assessment for Learning (AfL) and Assessment as Learning (AaL) (Zeng, et al., 2018) within the department. Respondents showed various levels of engagement with the Numbas system ranging from downloading pre-existing exams to creating new questions from scratch.

Overall results demonstrability was the most positive adoption attribute displayed by the respondents. This indicates that the degree to which the benefits of Numbas can be measured, observed and easily communicated to others is the main reason for the successful embeddedness of Numbas within the department. Voluntariness was also very positive but there were some more diverse opinions on this with some lecturers feeling that there was a requirement on them to use Numbas. The most interesting and unexpected result was that the lecturers disagreed that using Numbas would enhance their image or standing within the department. Image, at 2.11, was by far the lowest scoring attribute of Numbas. It also had the second lowest standard deviation at 1.15 indicating that there was a high level of agreement between participants. This means that they adopted the system in spite of this attitude being in contradiction of the findings of previous work. It is possible that this is in some way related to Irish culture as Ireland has a very low Power Difference Index, meaning that Irish people believe that power and status should be distributed equally (Hofstede, 1983). And so, it is perhaps not surprising that this group would find it difficult to see that using any one system would be a status boost as status is not easily seen or given.

The lecturers highlighted the need for training, both initial training to increase awareness of Numbas and ongoing or more advanced training to improve the skills of the adopters. In relation to training there was also comment about collaboration between new users of Numbas and the more experienced adopters as a form of ongoing training. There was mention that the use of Numbas should be broadened to more lecturers and courses and that purposes of use should also be extended e.g. to low stakes or no stakes homework. Lecturers commented on the time required to implement Numbas and cited the need for larger CIT specific question banks and additional funding as possible support. There were some concerns regarding the need to be fair to students in the marking of questions and also there was concern about the current inability to see all data relating to the question that an individual student was presented with an also concerns regarding loss of data between Numbas and the virtual learning environment. The final theme that was mentioned related to concerns about the pedagogical foundations of using Numbas and a need to see evidence of the teaching and learning benefits particular to Numbas that might differentiate it from other tools or methods.

7. Conclusions and Future Work

Numbas e-assessment has been used in Cork Institute of Technology for several years and the aim of this case study was to investigate whether Numbas is embedded within the department, how lecturers feel about Numbas and what lessons can be learned regarding successful implementation of Mathematics e-assessment. Diffusion of an innovation relies on many individuals deciding to adopt

the innovation and this depends on eight main attitudes of the individual towards the innovation: Voluntariness, Relative Advantage, Compatibility, Image, Ease of Use, Results Demonstrability, Visibility and Trialability. The case study was designed around an online questionnaire collecting mainly quantitative data with an element of embedded qualitative data. Quantitative and qualitative analysis was conducted on the data. There is broad use of Numbas within the department for various types of assessment and in various contexts, both inside and outside contact time. As expected the lecturers expressed positive perceptions towards most of the attributes with the notable exception of Image which runs contrary to the work of previous researchers. Themes emerging from the qualitative data included the need for ongoing staff training, concerns about fairness and transparency to students and ideas for the expansion of Numbas use. Future work will involve follow up interviews with some of the survey participants and, possibly more importantly, with some of the non-adopters who did not complete the survey.

8. Appendix

Survey Instrument adapted from Moore and Benbasat (1991) with the breakdown of the results for each item.

Attribute	Mean	SD
<i>Voluntariness (agree)</i>	5.5	2.26
My institute does not require me to use Numbas	4.89	2.71
Although it might be helpful, using Numbas is certainly not compulsory in my job.	6.11	1.62
<i>Relative advantage (somewhat agree)</i>	4.64	1.21
Using Numbas improves the quality of my work	4.78	1.39
Using Numbas enables me to accomplish tasks more quickly	5	1.18
Using Numbas makes it easier to do my job	5.22	0.97
Using Numbas enhances my effectiveness on the job	4.33	0.87
Using Numbas gives me greater control over my work	3.89	1.36
<i>Compatibility (somewhat agree)</i>	4.92	1.66
Using Numbas is compatible with all aspects of my work	4.56	1.74
Using Numbas fits into my work style	5.11	1.69
I think that using Numbas fits well with the way I like to work	5.11	1.69
<i>Image (disagree)</i>	2.11	1.15
People in my department who use Numbas have more prestige than those who do not	4.09	1.15

People in my department who use Numbas have a high profile	2.44	0.88
Using Numbas is a status symbol in my department	1.56	0.53
Ease of Use (somewhat agree)	4.85	0.88
I believe it is easy to get Numbas to do what I want it to do	4.44	0.73
Overall, I believe Numbas is (would be) easy to use	5.13	0.83
Learning to operate Numbas is easy for me	5	1
Results Demonstrability (agree)	5.81	1.21
I would have no difficulty telling others the results of using Numbas	6.33	1.32
The results of using Numbas are apparent to me	4.78	0.83
I would NOT have difficulty explaining why using Numbas may or may not be beneficial	6.33	0.71
Visibility (somewhat agree)	4.69	1.31
In my department, one sees many people using Numbas	4.33	1.12
Numbas IS very visible in my department	5.22	1.39
Trialability (neutral)	4.37	2.17
I am able to experiment with Numbas as necessary	4.89	1.76
Before deciding whether to use Numbas I was able to try it out properly	4.56	2.29
I was permitted to use Numbas on a trial basis long enough to see what it could do.	3.67	2.45

9. References

- Biggs, J. & Tang, C., 2007. *Teaching for Quality Learning at University*. 3rd Edition ed. s.l.:Open University Press.
- Carroll, T. et al., 2017. Numbas as an engagement tool for first-year Business Studies students. *MSOR Connections*, 15(2), pp. 42-50.
- Creswell, J. W., 2009. *Research Design: Qualitative, Quantitative and Mixed Methods Approaches*. 3rd Edition ed. s.l.:SAGE Publications.
- Foster, B., Perfect, C. & Youd, A., 2012. A completely client-side approach to e-assessment and e-learning of mathematics and statistics. *International Journal of e-Assessment*, 2(2).

- Hofstede, G., 1983. The cultural relativity of organizational practices and theories. *Journal of International Business Studies*, 14(2), pp. 75-89.
- Jenkins, M., 2004. Unfulfilled Promise: formative assessment using computer-aided assessment. *Learning and Teaching in Higher Education*.
- Loots, I., Fakir, R. & Roux, A., n.d. *An evaluation of the challenges and benefits of e-assessment implementation in developing countries: a case study from the University of Pretoria*. Cape Town, s.n., pp. 418 -420.
- McCann, A. L., 2010. Factors affecting the adoption of an e-assessment system. *Assessment & Evaluation in Higher Education*, 35(7).
- Moore, G. C. & Benbasat, I., 1991. Development of an Instrument to Measure the Perceptions of Adopting an Information Technology Innovation. *Information Systems Research*, September, 2(3), pp. 192-222.
- Rogers, E., 1962. *Diffusion of Innovations*. New York: Free Press.
- Sim, G., Holifield, P. & Brown, M., 2004. Implementation of computer assisted assessment: lessons from the literature. *Research in Learning Technology*, 12(3), pp. 215 - 229.
- Warburton, B., 2009. Quick win or slow burn: modelling UK HE CAA uptake. *Assessment & Evaluation in Higher Education*, 34(3), pp. 257 - 272.
- Zeng, W., Huang, F., Yu, L. & Chen, S., 2018. Towards a learning-oriented assessment to improve students' learning-a critical review of the literature. *Educational Assessment, Evaluation and Accountability*, August, 30(3), pp. 211 - 250.