The benefits and challenges of large-scale deployment of electronic voting systems: University student views from across different subject groups

Abstract

Electronic Voting System (EVS) is a classroom technology that provides a means to increase students' engagement, attention and attendance. The purpose of this paper is to provide a deeper insight into students' views on the benefits and challenges of EVS in the context of a large-scale institutional deployment and across different subject areas in higher education.

The data were collected from an online survey of 590 students across eleven academic schools at a UK university.

The non-linear principal component analysis of 32 question items from the survey showed that learning benefits, classroom-related benefits, usability and student-centered challenges are four distinctive dimensions in students' perceptions of the use of EVS. The non-parametric group comparison tests suggested that there are significant differences in learning benefits and challenges across different subject groups. However, the disparity appears to be related more to the way the EVS was used and the experience of students with it, rather than resulting from disciplinary differences. Content analysis of open questions revealed that summative use and staff competencies are the main issues related to EVS use by students. Finally, despite the overwhelming perception of the ease of use, it was found that usability could be an issue for students when EVS was used for summative assessment.

The implications of the study are: for practitioners, it underlines the importance of the focus on formative benefits of EVS as only then and regardless of disciplinary differences, can the promised rewards of the technology be gained; for institutions, it outlines some of the new challenges specific to the large-scale institutional implementation, judged through the lens of students' experience; for researchers, it provides an overview of the literature on large-scale deployment of EVS and it suggests some new areas for research on the use of EVS in higher education.

Keywords: improving classroom teaching; interactive learning environments; pedagogical issues; post-secondary education;

1. Introduction

Since their introduction in the late sixties (Kay & LeSage, 2009:819) Electronic Voting Systems, also known *inter alia* as "clickers", Audio Response Systems, or Personal Response Systems are being increasingly used in university education as a means to increase students' engagement, particularly in large classes. This is commonly achieved via in-class quizzes based on multiple-choice questions (MCQ), which can offer formative or summative outcomes. The technology behind EVS is very simple, consisting of a receiver, used for collecting the question responses and a number of handsets used for sending the responses by individual users. The examples of use cited in the literature are largely focused on small-scale implementation, within a single module and typically within subjects such as science, technology, engineering and mathematics (STEM). Using a mixed-methods approach, this paper seeks to address the gap in the literature and to explore the large-scale institutional implementation of EVS and differences in use and outcomes across academic disciplines, including those outside the STEM subjects.

1.1. Background

The research presented in this paper is part of a wider evaluation project initiated in the 2011/12 academic year that examined benefits and issues of large-scale institutional deployment of EVS technology for assessment and feedback at a medium-size post-92 UK university (UH) with in excess of 27,000 students. The aim of the project was to extend the benefits of EVS experienced by the early adopters from the area of Health and Human Sciences, across the whole university, while increasing the efficiencies related to the management of the handsets and the staff workload. The latter was achieved through an increase in the use of EVS for summative assessment. The central implementation campaign included regular staff training, installation of receivers in all teaching rooms, purchase of additional handsets, and the standardisation of operational procedures including distribution, ownership, and cost of handsets. By September 2012, 7265 EVS handsets had been purchased, 35% of the classrooms had been equipped with the EVS technology and 114 additional staff members i.e. 9.6% of the total staff population had been trained on technical and pedagogical use of the EVS.

The aim of the evaluation project, which worked alongside the institutional deployment project, was to support a scholarly approach to the changes in practice, while the purpose of this paper is to unravel some of the tacit and less explicit findings supported by the further analysis of the existing data. It is important to emphasise that although the data used for this paper are a subset of data collected during the evaluation project, they were collected purposefully in order to answer specific research questions, in an attempt to address some of the gaps identified in the current literature.

2. Literature review

The literature on EVS in university education dates from the 1990s (Dufresne et al., 1996; Mazur & Hilborn, 1997) to the present day and includes seven literature reviews conducted between 2002 and 2014. Amongst these, the review paper by Kay and LeSage's (2009) remains the most systematic and comprehensive appraisal of the literature, which has provided a framework for research in this area.

2.1. Benefits and challenges of EVS use

The review considered 64 peer-reviewed journal publications in the period 2000-2007 and found that the literature is pre-dominantly focused on learning and attitudes of undergraduate students studying STEM subjects, within a single module and in relatively large classes. In that context the following benefits were commonly identified:

- Classroom benefits: improvements in attendance, participation, and engagement;
- Learning benefits: the increased quantity and quality of class discussion; learning performance; quality of learning and contingent teaching;
- Assessment benefits: improved feedback; effective formative assessment; the ability to compare performance with others.

The review also highlighted some of the frequently reported challenges in EVS use:

- Technological challenges: missing or non-functioning handsets;
- Student-centered challenges: resistance to the new, active and more demanding teaching methods; the increased effort required from students when using EVS; some disruptions and confusion resulting from discussions; negative feelings towards the use of EVS for monitoring purposes, such as attendance and summative assessment.

The focus of subsequent research was on addressing specific gaps in the literature identified in the review and on providing empirical evidence for the findings from the review.

Thus, for example, Yeh and Tao (2013) found that student satisfaction is positively influenced by the learning- and classroom-related benefits; whereas it is negatively influenced by technology- and student-based challenges. Their results also suggest that assessment benefits do not have significant influence on students' satisfaction. One possible interpretation of the latter finding is that the students associated EVS assessment more with the negative aspects of tests, and less with the formative learning outcomes.

A recent study by Han and Finkelstein (2013) suggests that staff competencies and formative use of EVS are important indicators for students' perceptions of the EVS impact on their learning and engagement.

The findings from these papers, as discussed by the authors, could be specific to the context (a Taiwanese university in the former case and a novel research approach in the latter case) and therefore need further investigation.

Satisfaction with the technology, as discussed by Yeh and Tao (2013) is closely related to its usage, and can be considered in the context of TAM (Technology Acceptance Model) literature which postulates the "effort expectancy" and "performance expectancy" to be the most important determinants of intended and actual usage of the technology (Venkatesh et al., 2003). The "effort expectancy" related to the "ease of use" or "usability" of EVS has not been studied specifically in the EVS literature, as the underlying assumption is that the EVS technology is very easy to use. In the context of educational research, "performance expectancy" can be interpreted as learning efficiency gains and the gains in the final grade, neither of which has been explicitly mentioned in the EVS research.

The most recent literature review of the use of EVS in higher education by Han (2014a) considers the journal papers between 2010 and 2013, and focuses on finding the missing links across multiple EVS studies, within the intersection of the factors related to characteristics of students, characteristics of instructors, cognitive outcomes, non-cognitive outcomes and learning and teaching activities. The review confirms the findings from Kay and LeSage(2009), and it adds some new findings related to the impact of gender and previous achievement levels to students' learning with EVS. The most significant gaps in the literature are found in the area of the impact of students' and instructors' characteristics on the benefits of using EVS, as well as the relationship between levels and types of activities to the students' engagement and learning. The review also emphases the new methodological issues related to the lack of reporting on effect sizes, reliability test results, testing assumptions for use of parametric tests etc. The review however does not consider the technology factors such as usability and accessibility, or the environmental factors such as the readiness and equipment of the teaching rooms, and operational procedures related to the ownership, maintenance and distribution of the handsets, which are all important to consider in the context of institutional deployment of EVS.

2.2. Challenges of the large-scale institutional deployment of EVS

A significant increase in a scale of organisational adoption of EVS after 2002 has enabled the emergence of new institutional challenges such as handset distribution and management (Simpson and Oliver, 2007). While these are important to consider in the large-scale implementation, less is known about the students' view on benefits and challenges of EVS in this context, and how these compare to the findings from the research on small-scale implementation.

The findings reported in this section are based on a semi-systematic Google Scholar search of publications within the major academic databases including ACM Digital Library, EBSCO. Eduserv, Emerald, IEEE, ProQuest, JSTOR, Scopus, ScienceDirect etc.[i] The search was performed on titles and abstracts, containing one of following keywords: ARS, clicker, EVS, electronic voting, PRS, response system, in combination with the words learning, assessment, or feedback. Further selection was performed according to the following criteria: (1) only journal articles published between 2003 and 2015 are considered (2) the paper considered large (higher education) institutional deployment across multiple subject areas or multiple institutions.

The search and selection resulted in six publications, of which five were empirical studies (Draper and Brown, 2004; Kaleta and Joosten, 2007; Laxman , 2011; Yeh and Tao, 2013; Han and Finkelstein, 2013) and one (Twetten et al. 2007) was a practitioner report. While these studies vary in the scale of deployment and the size of survey samples (from a few hundred to over 5000 participants) they confirm the findings from the Kay and LeSage (2009) on the benefits and challenges of the small-scale use of EVS. In addition, they also reveal some of the new student-centered challenges such as:

- The importance of "staff competencies" for students' attitudes towards the EVS (Kaleta and Joosten, 2007; Han and Finkelstein, 2013);
- The importance of maintaining a focus on "pedagogy and learning" instead of on technology, as had sometimes been the case with overly enthusiastic adopters (Draper and Brown, 2004); similarly focusing on "pedagogy and learning" rather than on summative use (Han and Finkelstein, 2013) and related efficiency gains such as, administering exams and reducing the marking time for staff (Twetten et al., 2007);
- The expectations that EVS should provide "value for money" and be used more often if students pay for it (Kaleta and Joosten, 2007);
- EVS could be seen as an "additional burden" incurring additional cost, and the need to carry the handset around, to register the handset with MLE, to change the battery etc. (Twetten et al., 2007).

Specific to the scale and duration of deployment are the findings by Draper and Brown (2004) where they concluded that the perceived benefits of EVS increase through successive years of use.

Some of the studies exhibited certain limitations such as, lack of detail regarding the nature of deployment of EVS (Kaleta and Joosten, 2007, Yeh and Tao, 2013); occasionally employing informal and anecdotal evidence (Laxman, 2011; Twetten et al., 2007); a small number of items in the questionnaire (Han and Finkelstein, 2013).

2.3. Non-STEM subjects

The common intention of the above studies was to focus on the authors' use of EVS in their own STEM-based disciplines with the result that certain discipline areas such as social sciences have been severely under-represented in the EVS literature. Conclusions drawn from the scientific community cannot adequately report the use of EVS in more discursive subjects.

Draper and Brown (2004) have provided limited comparison across different subject areas; from the 9 disciplines they considered the only one outside of STEM was Philosophy. Their results suggest that the perceived benefits of EVS are relatively independent of the subject matter, but in fact Philosophy students were the least likely to state that they "definitely benefited" and most likely to be neutral or negative amongst all the disciplines.

Despite the call from Kay and LeSage (2009) to broaden the EVS research to non-STEM areas, reports on use of EVS outside of these subjects remain few, with the exceptions of

Business Marketing (e.g. Micheletto, 2011; Muncy and Eastman, 2012), Law (e.g. Easton, 2009; Steventon, Panesar, and Wood; 2012), and English (e.g. Cutrim, 2008; Cardillo, 2008). This was further confirmed by checking Bruff's (2014) bibliography of "clicker" studies which shows that the ratio of publications reporting the usage of EVS in STEM and non-STEM subjects is in the region of 4:1. While the aforementioned papers confirm some of the benefits and challenges identified earlier, they do not contribute any new subject-specific outcomes nor do they describe any new ways of using EVS, in comparison to the STEM subjects. Related to the latter, an exception is the study by Micheletto (2011), which reports on the use of EVS in discussions on sensitive and controversial topics such as business ethics, and suggests that the anonymity of EVS has enabled the collection of confidential data regarding self-reported behaviors and facilitated significant student reflection on perceived ethicality.

Another more recent study by Han (2014b) suggests that academic discipline and other contextual factors, such as, level of study are not significantly related to students' perceptions of learning with EVS. However the authors acknowledged that this might be due to some methodological constraints introduced in the study such as grouping of the subject areas (and study levels) into only two categories.

2.4. Summative assessment

The findings from Kay and LeSage (2009) suggest that although EVS is known to be used for assessing students' performance with formal grades, little research has been done to assess the impact it has on students' learning experience. This is confirmed in the studies by Kaleta and Joosten (2007) and Twetten et al. (2007), where EVS was used for "low-stake" assessment intended to stimulate discussion and assess students' comprehension. Han and Finkelstein (2013) showed that formative-only use of EVS has a more positive impact on students' perceptions on learning and engagement with the EVS, compared to other uses. Similarly as before, the type of use considered was either summative or formative, and it did not include combined use.

A study by Hancock (2010) reports on two trials, where EVS was used in a combination of formative and summative activities. The results of the first trial suggested that students' positive perceptions of formative EVS use shrank by 43% when it started being used as a replacement for traditional testing. In the second trial, the differences between formative and summative activities were blurred and EVS was used regularly in a formative way but with the summative outcomes contributing to 80% of the students' final grade. Despite the increased weight of the EVS questions in the final grade the satisfaction with the use of EVS rose by 8% compared to formative-only use. The results of this study imply the importance of careful and gradual introduction of summative EVS activities. Moreover, they suggest that even in summative tests, focus should remain on the formative benefits of the EVS. The main limitation of the study is the scope, as it considers a single subject area. Therefore, further research is required on the impact of different type of EVS use, including combination of formative and summative activities, across different subject areas.

3. Research questions

This research accordingly addresses some of the issues and gaps reported in the literature review above and considers the following research questions:

1. What are the main components contributing to the students' views on EVS in a context of a large-scale institutional deployment?
2. What are the new student-centered challenges that arise in a large-scale institutional

deployment of EVS?
3. What is the impact of increased summative use of EVS on students' views?
4. What are the differences in usage and student views' across different subject areas?

4. Methodology

This research is based primarily on the gathering of data through a purpose-designed student survey (*N*=590) with some additional comments provided by students studying at UH in 2011/12 via their online reflections. In the latter case, the students (*N*=26) captured their experiences of using EVS in class by either a daily written blog or by regular audio or webcams recordings. For the purpose of this paper these comments are used to illustrate the findings from the survey in the related discussion. Further details about the participants, structure of the online reflections and the related coding process can be found in the EVS project report.

4.1. Survey

The student questionnaire was made available online at the end of the spring term in 2012 and widely advertised to encourage participation. A prize draw for Amazon vouchers was offered as an incentive for students to participate. The average time for completing the questionnaire was estimated to be 15 minutes.

4.1.1. Questionnaire

The questionnaire comprised 42 closed questions and one open question, asking students to explain the reasons for not being satisfied with EVS use.

Amongst the closed questions, 6 were related to demographic information, 4 to the type of use of EVS and the remaining 32 were numerical items assessing students' views on EVS use. These items were using five-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree).

The usability of EVS was evaluated using the Simple Usability Scale (SUS), a research instrument developed by Brooke (1986). The scale consists of 10 questions and the method produces a single score between 0 and 100 to indicate the usability of the product. This study uses the modified version of the SUS questionnaire (questions labeled with * in Table 5) based on recommendations from Bangor, Kortum, and Miller (2008). The scale was augmented with an additional question on accessibility of EVS, which is considered to be an important indicator of the overall usability.

Four questions (labeled with ** in Table 5) adapted from Sun et al. (2008) were added to assess different levels of satisfaction, as per recommendations from the literature (Manisera, Van der Kooij, and Dusseldorp, 2010) which suggest that measuring satisfaction is complex, and requires multiple items to express the different aspects of satisfaction in a questionnaire.

The remaining seventeen numerical items (questions with no asterisks in Table 5) related to students' views on quality of learning and assessment, classroom related benefits, and performance expectancy were questions adapted from Draper and Brown (2004), Kay and LeSage (2009) and Venkatesh et al. (2003) respectively.

According to the guidance from DeVellis (2003) in order to avoid agreement bias the survey included a number of negatively worded questions (Table 5).

The Cronbach's α was calculated for all three subscales resulting in values: 0.85, 0.89 and 0.94 for usability, learning and satisfaction respectively, which were all above the minimal recommended value of 0.7.

4.1.2. Participants

The total number of participants was 590, a response rate of 14.41%, relative to the total number of 4093 handsets issued to students that academic year.

The student participants were self-selected but represented a wide cross-section of the UH student population as shown in Table 1. While the sample included more UG (undergraduate) students and fewer PG (postgraduate) students compared to the overall population, this was expected as the EVS deployment project was focused primarily on the UG courses.

Demographic		**Survey participants**		**UH student population**
Item	**Value**	***N***	**%**	**%**
Age	Less than 20	329	55.76%	48.81%
	21-24	172	29.15%	20.49%
	25 and more	89	15.08%	30.70%
Gender	Female	324	54.92%	56.08%
	Male	266	45.08%	43.92%
Nationality	UK	491	83.22%	81.34%
	International	99	16.78%	18.66%
Level	UG	571	96.78%	79.47%
	PG	19	3.22%	20.53%

Table 1 Comparison of demographic data between the survey sample (*N*=590) and the total UH student population

The majority of UG responses were from first year students (n=291, 49.32%). When compared with other attitude surveys (e.g. Han and Finkelstein, 2013) this year group was shown to be more likely to participate than other groups. Second year and final year UG students contributed to n= 136 (23.05%) and n= 144 (24.41%) responses respectively.

The experience of students with EVS was varied. The majority of students (n=288, 65.80%) had used EVS only in their current academic year, while 174 students (29.50%) had experience of using EVS in two consecutive academic years. A small number of students (n=28, 4.70%) had used them only in the previous academic year.

The distribution of responses across 11 different schools is displayed in Table 2 and it broadly matches the size of deployment of EVS in specific schools in the same academic year measured through a number of handsets purchased by the school at the time of the survey.

School	**Survey participants**		**Purchased EVS handsets**	
	N_1	%	N_2	%
Life Sciences:	99	16.80%	*1640	*22.57%
Law:	96	16.30%	1350	18.58%
Business:	94	15.90%	1000	13.76%
Computer Science:	74	12.50%	620	8.53%
Engineering and Technology:	63	10.70%	560	7.71%
Psychology:	56	9.50%	*	*
Health and Emergency Professions:	46	7.80%	600	8.26%
Education:	38	6.40%	500	6.88%

Humanities:	11	1.90%	675	9.29%
Physics, Astronomy, Mathematics	8	1.40%	320	4.40%
Creative Arts:	**5	0.80%	**0	0%

Table 2 Distribution of student responses (N_1=590) and handsets purchased by individual schools by the end of 2011/12 (N_2=7265).
*The figures for Psychology handsets have been reported together with Life Sciences, according to the school divisions prior to 2011.
**A small number of handsets (70) were available for hiring from the central office.

4.1.3. Subject groups

In order to look at the broad subject-specific differences, the following three subject groups are considered:

- Health and Human Sciences (HHS), comprising the schools of Life Sciences, Psychology and Health/Emergency Professions;
- Social Sciences, Arts and Humanities (SSAH), comprising the schools of Law, Business, Education, Humanities and Creative Arts;
- Science and Technology (ST), comprising the schools of Computer Science, Engineering and Technology, and Physics, Astronomy and Mathematics.

Subject group	**Survey participants**		**UH student population**	**Purchased EVS H/S**
	N	%	**%**	**%**
HHS	201	34.07%	32.46%	30.83%
SSAH	244	41.35%	52.97%	48.52%
ST	145	24.58%	14.57%	20.65%

Table 3 Counts and percentages (*N*=590) of survey responses, student population, and number of EVS handsets (H/S) purchased in 2011/12 across the subject groups

The distribution of survey responses across the three subject areas, broadly matched the size of these groups in the total UH student population, as well as the size of EVS deployment in these groups, measured through the number of handsets purchased by the groups at the time of the survey (Table 3).

4.1.4. Procedures/EVS Usage

The type, frequency and duration of EVS use varied.

As shown in Table 4, the most frequent type of use was a combination of summative and formative activities, followed by summative-only or formative-only use. For a small percentage of students, EVS was used only for other non-teaching-related purposes such as research, module feedback, opinion polls or for attendance monitoring.

Type of use	***N***	**%**
Summative and formative	221	37.46%
Summative (only)	178	30.17%
Formative (only)	118	20.00%
Summative and formative and attendance	23	3.90%
Summative and attendance	17	2.88%
Other mainly non-teaching use	14	2.37%
Attendance (only)	5	0.85%
Formative and attendance	5	0.85%

Summative and formative and other mainly non-teaching use	5	0.85%
Formative and other mainly non-teaching use	2	0.34%
Summative and other mainly non-teaching use	2	0.34%

Table 4 Distribution of the counts and percentages (*N*=590) of student responses related to the type of EVS use

Regarding the scope, frequency and duration of use, a majority of students had used EVS in 1-3 modules (n=371, 62.88%), once a week (n=349, 59.20%), and for up to 10 minutes per session (n=404, 68.50%).

5. Data Analysis and Results

The first part of this section includes the rationale for using the specific data analysis techniques in relation to the research questions under the consideration.

Numerical items were analysed with SPSS v19.0.0 using several different techniques, outlined below.

To answer the first research question (“What are the main components contributing to the students’ views on EVS in a context of a large-scale institutional deployment?”) the Principal Component Analysis (PCA) technique was used (Section 5.1).

PCA is a technique used for reducing a number of variables into a smaller set of “principal components” representing significantly correlated clusters of variables, while retaining as much of the original information as possible. Recently, this technique has gained in popularity in research into educational technologies, where it is used to assess the consistency of the questionnaire subscales (e.g. Biasutti, 2011); or to reduce the dataset into a small set of components for further analysis with other variables (e.g. Han and Finkelstein, 2013). However, standard PCA is based on the assumptions that the variables are of numerical (interval or ratio) measurement level and that relationships between them are linear. Since none of the assumptions could be justified for the type of data considered in this paper (Likert-scale variables) the non-linear PCA (Manisera et al., 2010) was applied instead. More specifically, the CATPCA (CATegorical PCA), the SPSS implementation of non-linear PCA was used to explore students’ views on the use of EVS, and to simplify group comparison against the fewer elements. The dataset considered for the CATPCA included 32 Likert-scale variables. As recommended by Manisera et al. (2010) the missing values, which counted for less than 2% of values, were imputed with the corresponding mode. Since the CATPCA does not offer rotation option, and because transformations are invariant under rotation, the transformed variables were used as input for a classical PCA with oblique rotation, in order to simplify the interpretation of the components (Manisera et al., 2010). In educational research a similar technique was employed by Gil-Flores, Torres-Gordillo and Perera-Rodríguez (2012) and Hamlen (2012), amongst others.

In section 5.1.1, the SUS score was calculated according to the instructions from Brooke (1986) and Bangor, Kortum, and Miller (2008), and this technique contributed to further understanding of the students’ views on the EVS use (see section 3 first research question).

Content analysis is a recommended technique for coding the open ended questions in surveys (Cohen, Manion and Morrison, 2000) and it has been used in this research to analyse the textual responses to the question asking students to explain the reasons for dissatisfaction with the EVS use (see section 3 second research question). Coding was performed with HyperResearch v.3.5.2. The unit of coding was a single sentence, and the coding process proceeded in three phases. In the first phase the source items were coded into the pre-existing categories found in the literature on small-scale (Kay and LeSage, 2009) and large-scale implementation (Section 2.2.). In the second phase, the items that could not be mapped into any of the existing categories were summarized and abstracted into the new codes, using the techniques described in Cohen et al. (2000). Finally, all codes were augmented with an additional qualifier wherever a repetitive sub-theme was encountered in responses. For example, the source item "*EVS has caused me stress as 10% of my final grade is based on it"* was coded as SummativeAssessment:EVS%TooHigh, where "SummativeAssessment" was the code corresponding to one of the "challenge" categories from the Kay and LeSage (2009) review, and "EVS%TooHigh" was an additional qualifier, corresponding to a sub-theme of "SummativeAssessment".

Kruskal-Wallis and Jonckheere-Terpstra tests were employed to reveal the impact of the type of use and the subject area to those aspects of students' views discovered by the PCA, thus corresponding to the third and fourth research questions respectively. These tests compare the medians of the groups, and in the latter case, test for an ordered pattern of the medians. The tests are adequate for comparing more than 2 groups when parametric assumptions are not met. For the type of data considered here, the CATPCA transformations produce the values which respect the rank order, but do not necessarily have equal intervals, which justified the use of non-parametric tests.

The results of all statistical tests used in the paper are reported using the guidelines provided in Field (2013).

As recommended by Han (2014a) the results of all statistic tests are reported with the corresponding effect sizes; the results of reliability tests are provided where required and whenever the data considered are of the interval type, the corresponding non-parametric tests are used.

The rest of this section includes the presentation of the results of the survey data analysis.

5.1. First research question

An initial PCA was conducted first on all 32 numerical items considered in this paper, resulting in the removal of one variable ("Lecturers are changing their style of lectures because of EVS") because of a low correlation with other items and the vagueness of the question. CATPCA was then conducted on the remaining 31 items, resulting in 6 components with the eigenvalues higher than 1. The scree plot was slightly ambiguous and showed inflexions that would justify retaining both components 2 or 4. Given the large sample size, the convergence of the scree plot and the small contribution of the component 5 and 6 to the total variance explained (less than 8%), four components were retained in the final analysis. The rotated PCA (oblimin) was then performed on the transformed variables. Four components in combination explained 39.57%, 10.34%, 6.07%, and 4.89% of variance respectively and 60.86% in total, which is comparable to other studies which adopted the similar measures (e.g. Manisera et al., 2010; Gil-Flores et al., 2012).

All validation tests required for the PCA confirmed the sampling adequacy and the fit of the obtained model: the average communality 0.61 exceeded 0.6, the recommended value for samples of size greater than 250; the Kaiser-Mayer-Olkin (KMO) measure = 0.95, and all KMO values for individual items were >0.89, above the acceptable limit of 0.5; Bartlett's test of sphericity $\chi^2(465)= 12226.43$, $p< 0.001$, indicated that correlations between items were sufficiently large for PCA; the reproduced correlation matrix showed 129 (27.0%) non-redundant residuals with absolute values greater than 0.05, which is below the necessary 50% mark.

Table 5 shows the components' loadings after rotation. The results were confirmed with the corresponding structure matrix (Appendix A), and with an additional PCA performed for each component, showing that all components are one-dimensional, with Cronbach's α 0.95, 0.87, 0.80, and 0.86 respectively.

The items that cluster together suggest that components 1-4 represent respectively:

1. Learning benefits (quality of learning, feedback)
2. Student-centered and technology-related challenges (staff competencies, ease of use, technical support, use for testing)
3. Ease of use of the EVS technology
4. Classroom-related benefits (engagement, participation, and motivation).

The component loadings for the four satisfaction items (labeled with ** in Table 5) show that they are positively related to components 1,3 and 4 and negatively related to component 2. With regards to the questionnaire sub-scales, all items from the "satisfaction" scale (** in Table 5) mapped into the "learning benefits" component; positive questions from the "usability" scale (* in Table 5) mapped onto the "ease of use" component, while negatively phrased "usability" questions mapped onto the challenges (second component). The "learning and assessment" scale (all other items in Table 5) split into "learning benefits" and "classroom related benefits" components respectively.

Pattern Matrix[a]	Components			
	1	**2**	**3**	**4**
Using the EVS allows problem areas to be identified	**0.866**			
EVS provides an immediate check of students' understanding	**0.836**			
Viewing responses gives me an idea of how I am doing in relation to others	**0.799**			
Responding to questions in class makes me think about the course	**0.789**			
The lecturer addressed relevant topics or issues identified in responses	**0.687**			
I think I would like to use EVS frequently *	**0.519**			
Using EVS enables me to accomplish learning activities more quickly	**0.513**			0.442
The advantages of using EVS outweigh any disadvantages **	**0.513**	-0.167	0.105	0.278
I enjoy using EVS in my learning **	**0.484**	-0.147	0.12	0.415
I look forward to using EVS in the future **	**0.46**	-0.177	0.155	0.446
I am very satisfied with class sessions that use EVS **	**0.428**	-0.156	0.162	0.406
Sometimes it is not clear what I am supposed to be voting for		**0.771**		
I found EVS very awkward to use *		**0.713**		
I needed to learn a lot of things before I could get going with EVS *		**0.69**		
I think that I would need the support of a technical person to use EVS *		**0.668**		
Using EVS can distract from the learning point entirely		**0.631**		
Setting up and using the handsets takes too much time in lectures		**0.598**		
Sometimes lecturers seem to be asking questions just for the sake of it		**0.592**		
I thought that there was too much inconsistency in the EVS functions *		**0.584**		
I found EVS unnecessarily complex *		**0.539**	-0.403	
Often, I do not have enough time to think before I have to vote		**0.486**		
I thought EVS was easy to use *			**0.675**	
I would imagine that most people would learn to use EVS very quickly *			**0.613**	
I felt very confident using EVS *			**0.587**	
The EVS handset is suitable for a student with disabilities			**0.567**	
I found that the various functions in EVS were well integrated *			**0.442**	
I ask more questions in classes that use EVS				**0.874**
I think about the subject outside of the class more in classes that use EVS				**0.87**
I spend more time discussing the subject in classes that use EVS				**0.842**
I am motivated to attend classes that use EVS				**0.793**
Using EVS will increase my chances of getting a good grade				**0.391**

Table 5 Results of CATPCA followed by rotated PCA (components highlighted in **boldface**)
a. Rotation converged in 23 iterations.* SUS items **Satisfaction items

5.2.1 SUS Score Calculation

The SUS score for EVS was calculated to be 74%. According to Bangor et al. (2008) benchmarking, this score places EVS above "good" (M=72.75) but below "excellent" (M=85.58) in the "Mean SUS table". That makes EVS more usable than technologies such as mobile phones (in use prior to 2008), web pages and interactive voice response systems, but less usable than graphical user interfaces. A small, but not insignificant number of students (n=40, 6.78 %) disagreed that EVS is an accessible technology.

5.2. Second research question

A total of 121 (20.51%) student responses corresponding to 243 source items were coded into the 16 categories shown in Table 6.

Challenge	Source	*N*	%
Summative assessment	Lit. on small-scale impl.	58	23.87%
Staff competencies	Lit. on large-scale impl	49	20.16%
EVS did not work	Lit. on small-scale impl.	46	18.93%
Focus on pedagogy and learning	Lit. on large-scale impl	29	11.93%
Value for money	Lit. on large-scale impl	16	6.58%
Attendance for grades	Lit. on small-scale impl.	11	4.53%
New method	Lit. on small-scale impl.	11	4.53%
Inadequate for testing subject knowledge	NA	5	2.06%
Effort	Lit. on small-scale impl.	4	1.65%
Accessibility	Lit. on small-scale impl.	3	1.23%
Additional burden	Lit. on large-scale impl	3	1.23%
Bringing EVS	Lit. on small-scale impl.	3	1.23%
Discussion	Lit. on small-scale impl.	3	1.23%
Inconsistency of use	NA	2	0.82%
Identifying students	Lit. on small-scale impl.	0	0.00%
Negative feedback	Lit. on small-scale impl.	0	0.00%

Table 6 Counts and percentages (*N*=243) of challenges identified in the free text answers to the question "If you are not satisfied with using EVS please explain …". New challenges are shown in **boldface**.

The total number of items matching the challenges from the literature on small-scale implementation (Kay and LeSage, 2009) is n=139 (57.20%) while the challenges specific to the large-scale implementation (Section 2.2) contributed to n=97 (39.92 %) items. The remaining seven items (2.88 %) correspond to the "emergent" themes not previously reported in the literature, such as inadequacy of EVS for testing the subject knowledge ("*Being law students, many questions need to be debated as there may be several different answers which could apply in different circumstances. The EVS assessments do not allow us to do this")* and inconsistency of use ("*Lecturers are not consistent with the way in which we use EVS. Timed, not timed, open book, not open book").*

The subthemes and illustrative students' quotes for the most frequent categories (> 10%) from Table 6 are shown in Table 7 below.

(a) Summative Assess.	*N*	%	Quote (unedited)
Technology unreliable	24	9.88%	*The only issue I have is the worry that they could go wrong during an assessment and it affects your score. Is there any way there could be "spare" EVS in case say batteries run out, or technology fails!*
Cheating	17	7.00 %	*It was hard to conceal the response from others*

Not enough time to answer	7	2.88 %	*I do not like tests involving EVS as I feel there is not enough time to think about the question thus resulting in a rushed answer*
EVS% too high	7	2.88 %	*EVS has caused me stress as 10% of my final grade is based on it*
Cannot change the answer	3	1.23 %	*There is no opportunity to go back and change your mind*
(b) Staff Competencies	***N***	**%**	**Quote**
Technology	42	17.28%	*The lecturers could not always get the system to work first time, and that meant that the majority of the students digressed in to talking about subjects not related to the session.*
Quality of questions	7	2.88%	*In some modules the wording of the questions and answers is not particularly clear.*
(c) EVS did not work	***N***	**%**	**Quote**
EVS did not work	29	11.93%	*Sometimes the handsets do not work.*
Answers not recorded	13	5.35%	*Technical problems with receivers not picking up the devices and this takes up time in the lectures.*
Issues with channels	3	1.23%	*… me personally I find the EVS rather temperamental when trying to change channel for specific rooms.*
Signals overlapping	1	0.41%	*… when a neighbouring class simultaneously use the EVS, the digital receivers are able to pick-up their answers as well as ours*
(d) Focus on …	***N***	**%**	**Quote**
Assessment	11	4.53 %	*The lecturers did not elaborate from the results of the EVS, we were not given the chance to discuss them, they were calculated and then moved on from*
Factual knowledge	9	3.70%	*(the EVS) seems to only assess recognition memory rather than whether the student understands the topic or not.*
Technology	9	3.70%	*I would just be aware that sometimes the voting device can be seen as the most interesting item and the point of learning gets lost*

Table 7 Counts (total = 182) percentages (*N*=243) and illustrative student quotes for the sub-themes of the most frequent themes from Table 6.

5.3. Third research question

In order to simplify the analysis of the impact of increased use of EVS for summative assessment, the following three groups of students were considered,

1. Students who used EVS only for non-summative tests
2. Students who used EVS through a combination of summative and non-summative tests,
3. Students who used EVS only for summative assessment.

Table 8 shows the medians and interquartile ranges of the four principle components across the three groups.

The Kruskal-Wallis test revealed that "learning benefits" and "challenges" (1st and 2nd component in Table 5) were both significantly affected by the type of use (Table 8) where $H(2) = 27.23$, $p < 0.001$ and $H(2) = 10.09$, $p < 0.01$ respectively.

Mann-Whitney tests were used to follow up this finding. A Bonferroni correction was applied and so all effects are reported at 0.05/3=0.0167 level of significance, where 3 is the number of additional tests performed. The results indicate that,

- Learning benefits were significantly lower in the groups which used only summative tests compared to the other two groups: non-summative ($U = 10640$, $r = -0.15$) and combination ($U = 17014.5$, $r = -0.24$).
- Similarly, challenges were significantly higher in the groups which used only summative tests compared to the other two groups: non-summative ($U = 10418$, $r = -0.16$) and

combination (U=20325.5, r=-0.11).

Component	Type of Use	n	% (N=590)	Median	Interquartile Range	Mean Rank
1: Learning benefits	Non-summative	144	24.41%	-0.0026211	0.94296	292.44
	Combination	268	45.42%	0.2435986	1.00383	330.78
	Summative	178	30.17%	-0.0893559	1.22415	244.86
2: Challenges	Non-summative	144	24.41%	-0.2614623	1.17034	274.09
	Combination	268	45.42%	-0.1670862	1.24448	284.9
	Summative	178	30.17%	0.0319619	1.22785	328.78
3: Ease of use	Non-summative	144	24.41%	-0.0826177	1.56968	292.67
	Combination	268	45.42%	-0.0469942	1.51293	302.73
	Summative	178	30.17%	-0.1570742	1.53808	286.91
4: Engagement	Non-summative	144	24.41%	0.2245773	0.91904	288.51
	Combination	268	45.42%	0.2977497	0.62415	305.77
	Summative	178	30.17%	0.2677923	1.16127	285.69

Table 8 Medians and interquartile ranges for components 1-4 across the groups with different types of use

5.4 Fourth research question

The size of EVS use, the type of use and the students' experience with use, differed across the three subject areas, as shown in Figure 1. Chi-square test confirmed the existence of significant associations (all at p<0.001 level) between the subject area and the corresponding variables: (a) number of modules: $\chi^2(6)= 77.17$ (b) students' experience: $\chi^2(4)= 131.16$, and (c) level of summative use: $\chi^2(4)= 152.58$ respectively. Moreover, Cramer's V statistics for level of summative use (0.35) and students' experience (0.33) were both above the boundary for medium effects (0.3) suggesting that the strength of these relationships is important.

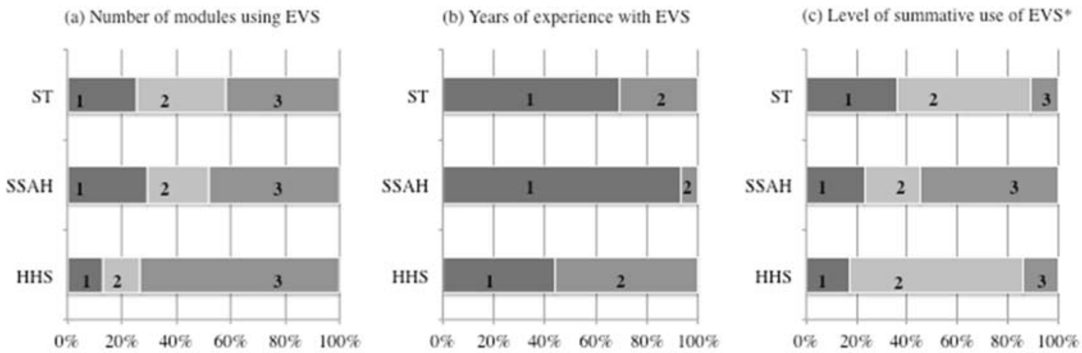

Figure 1 Differences in the EVS usage across subject groups.
*1=No summative use, 2=Combination of summative and non-summative use, 3=Only summative use.

Table 9 shows the medians and interquartile ranges of the four principle components across the three subject groups.

The Kruskal-Wallis test revealed that "learning benefits" and "challenges" (1st and 2nd component in Table 5) were both significantly affected by the subject group (Table 9) where H(2) =29.82, p< 0.001 and H(2)= 24.97, p<0.001 respectively.

Mann-Whitney tests were used to follow up this finding. A Bonferroni correction was applied and so all effects are reported at 0.05/3=0.0167 level of significance, where 3 is the number of additional tests performed. The results indicate that,

- HSS group reported significantly higher "learning benefits" compared to the other groups: SSAH (U= 17339.5, r=-0.25) and ST (U= 11531, r=-0.18).
- SSAH group reported significantly higher "challenges" compared to the other groups: HHS (U= 17859.5, r=-0.23) and ST (U= 15019, r=--0.12).

Component	Subject group	N	% (N=590)	Median	Interquartile Range
1: Learning benefits	HHS	201	34.07%	0.2820678	0.94718
	SSAH	244	41.36%	-0.0745569	1.09967
	ST	145	24.58%	-0.0032942	0.97743
2: Challenges	HHS	201	34.07%	-0.3067031	1.06542
	SSAH	244	41.36%	0.1489725	1.33369
	ST	145	24.58%	-0.1532863	1.15544
3:: Usability	HHS	201	34.07%	-0.1508076	1.35973
	SSAH	244	41.36%	-0.0759064	1.61548
	ST	145	24.58%	-0.0911381	1.66329
4:: Engagement	HHS	201	34.07%	0.2896112	0.66309
	SSAH	244	41.36%	0.2635801	1.13535
	ST	145	24.58%	0.2852685	0.66212

Table 9 Medians and interquartile ranges for components 1-4 across different subject groups

Given the significant difference between the groups with respect to the students' experience with the EVS use (Figure 1b), another Jonckheere-Terpstra test was performed to check for an ordered pattern in the medians of the groups that used EVS only in one academic year, compared to two consequent years (first two rows for each group in Table 10). The test results revealed a significant trend in the data: as the experience with EVS increased, "learning benefits" increased (J=43917.5, z= 4.09, r=-0.17, p<0.001), "challenges" decreased (J=30583.5, z= -2.97, r=-0.12, p<0.01), and "classroom benefits" increased (J=41043.5, z= 2.57, r=-0.11, p<0.001).

Component	Period of use	n	% (N=590)	Median	Interquartile Range
1: Learning benefits	this year only	388	65.76%	-0.0083189	0.89213
	this year and last year	174	29.49%	0.3028152	1.11007
	last year only	28	4.75%	0.0074359	1.05429
2: Challenges	this year only	388	65.76%	0.0011923	1.23376
	this year and last year	174	29.49%	-0.2967323	1.21655
	last year only	28	4.75%	-0.3151505	1.34506
3:: Usability	this year only	388	65.76%	-0.1537075	1.57671
	this year and last year	174	29.49%	-0.0327027	1.42129
	last year only	28	4.75%	0.3362816	1.54973
4:: Engagement	this year only	388	65.76%	0.2305356	0.90111
	this year and last year	174	29.49%	0.3437869	0.68371

last year only	28	4.75%	0.3134735	0.8817

Table 10 Medians and interquartile ranges for components 1-4 across the groups with different periods of use

6. Discussion

The goal of the paper is to offer a deeper insight into students' views on the benefits and challenges of EVS in the context of a large-scale institutional deployment and across different subject areas and different levels of summative and formative use. This section includes the discussion of the results and limitations of the study and it is structured according to the initial research questions.

6.1. What are the main components contributing to the students' views on EVS in the? context of a large-scale institutional deployment?

The results on the CATPCA indicate that there are four distinct components determining students' views on EVS use: learning benefits, classroom-related benefits, usability and student-centered challenges, covering a broader range of students' perceptions compared to similar studies (e.g. Han and Finkelstein, 2013).

The component loadings of the satisfaction items (** in Table 5) confirmed Yeh and Tao's (2013) findings on students' satisfaction being positively related to learning and classroom-related benefits and negatively related to technology- and student-centered challenges. Assessment benefits, such as an immediate check of understanding and the ability to compare with others were included in the "learning benefits" component, and as such had a positive relation with students' satisfaction, unlike in Yeh and Tao (2013) where no relation was found. The difference could be attributed to the context (Taiwanese vs. UK university) but also to the more discriminative wording of related questions in their questionnaire e.g. "Using PRS enables the assessment to be done to improve student understanding", compared to this research: "EVS provides immediate check of students' understanding".

Learning and classroom related benefits of EVS were confirmed in students' online reflections, and in some cases even further clarified:

- *"EVS encourages more thinking" (AJ8)*
- *"EVS does not help learning directly, but it helps break the lecture and makes it more interactive" (AJ2).*
- "*EVS is more helpful for people who would not engage otherwise*" (MC13).

Despite the EVS scoring high on the usability scale, students' views on the "ease of use" were divided. The spread of usability items (* in Table 5) across components 2 and 3, suggests, that although there is a significant albeit small correlation between usability component and satisfaction items (** in Table 5) students become increasingly critical of the usability of EVS when other problems related to staff competencies and the use for testing, are encountered. These results are providing empirical confirmation for some anecdotal evidence from the literature, such as that understanding the handset operations and "indicator lights" are not as straightforward as they seem (Twetten et al., 2007). The divided views on the usability of EVS were confirmed also in students' online reflections, some of which highlighted accessibility issues:

- *"It is very easy to use the EVS as all you have to do is press the button..."* (AJ2);
- *"I could use the basic features on it, but some buttons I feel afraid to touch, as I don't know what they do." (*MC1*).*
- "*It is a fairly small screen to look at and user may also hit the wrong button because I feel the buttons are really close together"* (AJ6*).*

Challenges encountered in the literature on the small-scale implementation (Kay and LeSage, 2009) contributed to more than half of all items identified in the student responses to the open question (Table 6) suggesting that the scale of implementation makes no difference with regards to the known issues with the EVS use. In other words, despite the increased investment in staff training and classroom equipment, students continue to report that EVS did not work, and continue to dislike the use of EVS for summative assessment. This suggests that there are other more profound factors in the students' dislike of using the technology for formal assessment. The research on students' perceptions of the technology-enhanced learning (TEL) points out that personal characteristics such as learner's technology anxiety and self-efficacy could significantly influence the satisfaction with TEL (Sun et al., 2008).

6.2. What are the new student-centered challenges that arise in a large-scale institutional deployment of EVS?

The results of the content analysis (Table 6) confirmed many of the challenges reported in the literature on large-scale implementation. The most frequent amongst these challenges are staff competencies on technical as well as pedagogical aspects of the EVS use. Related to the later, is the challenge of maintaining the focus firmly on pedagogy and learning, rather than technology and assessment, which was not always the case according to student responses (Table 7d).

These results are providing further empirical evidence for similar findings from the literature, such as that students are becoming increasingly critical if the staff member cannot resolve problems with the equipment quickly (Kaleta and Joosten, 2007) and also confirming early findings by Draper and Brown (2004) that success depended on putting pedagogy first, technology second.

The findings differ from the results reported by Han and Finkelstein (2013) as they were assessing the impact of the level of staff EVS training to students' perceptions on learning and engagement. At UH, all staff who used EVS in 2011/12, were provided with the same level of training. These sessions were designed to provide staff with the 'hands-on' experience with EVS, and to educate them on specific usage examples and related pedagogical benefits i.e. their aim was to decrease the "effort expectancy" and increase "performance expectancy" for staff, and subsequently and according to TAM theories, positively influence their attitudes towards the EVS use. Therefore, dissatisfaction of some UH students with the staff EVS competencies could not be attributed to the level of training, but possibly to other moderating factors, identified in the TAM literature such as age, gender, experience and voluntariness of use (Venkatesh et al., 2003).

The technology-centered challenges included some issues specific to the increased use of EVS for summative assessment, such as the generally un-substantiated fear that answers are not recorded properly. The large-scale deployment also resulted in occasional specific technology-related problems such as time spent trying to find a correct EVS channel due to signal interference between neighbouring classrooms' EVS systems. While these issues are likely to appear when the EVS is deployed for the first time on a large-scale, they should be easily resolved in subsequent implementation of the technology, as was the case at UH.

Emergent challenges from students' responses such as inadequacy for testing the subject knowledge were re-iterated in students' online reflections:

- *"It would be difficult to use it in 'Language, Law and Politics' for example as there is a lot more text and theoretical information rather than just one straight answer for a question."(MC1)*

Rather than inadequacy of the subject, this and similar views (Section 5.3) indicate that the issue might be with the experience of tutors from SSAH area in writing questions for testing higher-order knowledge domains in their disciplines.

Challenges related to “value for money” and “additional burden” were reported at a smaller extent (n=19, 3.22%) indicating that the operational procedures related to handset ownership, cost and distribution worked well, as noted in the final project report.

While it could be argued that some of the challenges specific to the large-scale case can be equally observed in a small, single-module implementation, the size of the deployment has created an environment where these problems became more visible. For example, the issues related to staff competencies and (the lack of the) focus on pedagogy are less probable in a single-module implementation, where the teaching staff are likely to be early adopters, enthusiastic about and competent with the technology and its affordances.

6.3. What is the impact of increased summative use of EVS on students’ views?

The results of group comparison (Section 5.4.) indicate that the use of EVS only for summative tests is related to lesser learning benefits and higher challenges for students. According to the survey responses (Table 6) the use of EVS for summative assessment has been the main reason for students’ dissatisfaction with the EVS. These results clearly indicate that the increased use of summative assessment across UH had an overall negative impact on students’ views of EVS.

These results provide further empirical evidence for the initial findings from Kay and LeSage (2009) and they also extend the results by Han and Finkelstein (2013) on the negative impact of summative assessment on students’ perceptions on learning and engagement. With respect to the latter, this research considers three different levels of summative use, rather than only two categories (Section 5.4.). This is important, as according to Hancock (2010) the combination of summative and formative use can lead to even higher satisfaction than in formative-only use.

The increased use of summative assessment has made the students more anxious as the technology was not perceived sufficiently reliable for summative use (Table 7a). This is similar to Kaleta and Joosten’s (2007) findings, that even with “low stake” summative assessment the smallest issues with the technology could become detrimental and create unnecessary anxiety for some students. One of the difficulties is to determine what constitutes the “low-stake” assessment, as for some students even 10% could be a reason to “stress” (Table 7a), while for others, 80% is acceptable and results in increased satisfaction compared to formative-only use (Hancock, 2010). These findings suggest that more important than the percentage, are the formative benefits that can be achieved in summative use. The benefits of formative use of EVS are not confined only to students’ satisfaction but they also extend to the learning performance. One example that supports the latter claim is the recent experimental study by Lantz and Stawiski (2014) where it was found that EVS questions improved memory for material two days later compared to no-EVS controls, provided that immediate feedback was given about each question.

Other common reasons for dissatisfaction with the summative use, such as increased opportunities for cheating, time to answer and inability to change the answer are discussed in the literature and suggestions for overcoming them are proposed (Caldwell, 2007).

Some of the online reflections confirmed the negative views of students towards the summative use of EVS, if the opportunity to discuss answers was not given:

- *“The tests that counted, tended to leave a certain bad feeling amongst some of the students; the tutors did not want to reveal the right answer, so as to give answers to* *groups undertaking the same test but later in the week"* (MC10).

6.4. What are the differences in usage and student views' across different subject areas?

The subject-specific differences in EVS usage and student views have been studied only by a few researchers, and no significant differences in perceived benefits of EVS across different subject areas were found (Draper and Brown 2004, Han and Finkelstein, 2014).

The results of the CATPCA component comparison undertaken here across three different subject areas (Section 5.5) indicate that there are significant differences with respect to students' perceptions of learning benefits and challenges.

However, a closer inspection of the details of the usage in the three subject groups suggests that the differences could be attributed to the level of summative use and students' experience with the EVS rather than any other disciplinary differences:

- In HHS and ST the EVS was used mainly through a combination of low-stake assessment and formative activities such as revision, spot-tests and debates, while in SSAH the use was dominated by the assessed in-class tests.
- A majority of HHS students have used EVS in two subsequent academic years, whereas in SSAH and ST the majority of students were the first time users.
- The size of use, measured through a number of modules where EVS was used and the frequency of use were both higher in HSS compared to the two other schools.

It is not surprising then that the HHS students, reported higher learning benefits than other groups, and similarly that SSAH students found the EVS use more challenging. The question is why was EVS used differently in different subject groups? Possible reasons are: the experience of staff in HHS, which was the group with the highest percentage of early adopters amongst staff; large class sizes in some of the SSAH subject areas might have influenced the staff perceptions of EVS as a tool for reducing the marking time; and voluntariness of use was not always achieved in one of the SSAH schools where EVS was used across all first year modules.

6.5. Limitations

Usability and classroom benefits (components 3 and 4 in Table 5) explained relatively small percentage of the variance in data, compared to the other two components. This is however comparable to results from other studies using CATPCA e.g. Manisera et al. (2010). Reducing the number of components was not perceived to be helpful for the interpretation of the results.

As expected, the components' correlation matrix (Appendix A) shows them to be significantly related ($p<0.01$). While other correlations are weak, the level of correlation between the learning benefits and classroom benefits is 0.53, which indicates a certain overlap in the latent variables measured by the two components. This is not surprising given that students' engagement is known to be one of the better predictors of learning (Carini, Kuh and Klein, 2006).

Some of the findings have emphasised the incompleteness of the survey instrument e.g. not taken into consideration were the personal characteristics of individual learners related to the technology self-efficacy, which could influence their attitudes towards the summative use of EVS.

The survey response rate was lower than expected, but the sample was a good representative of the overall UH population with regards to age, gender, and subject area. However it should be noted that the sampling was by self-selection, and that those participating are possibly the more engaged students and these may be more favourably disposed towards innovations introduced by tutors than those who are less engaged.

7. Conclusions

The benefits of EVS for students' learning, engagement and motivation are well-known from the numerous studies on small-scale implementation. The aim of this research was to obtain a deeper insight into students' views on the benefits and challenges of EVS in the context of a large-scale institutional deployment and across different subject areas, the topic that is currently under-represented in the EVS literature.

The results of this research suggest that the scale of EVS deployment can create an environment for new challenges to surface, especially those related to the use of technology by less experienced teachers and too much emphasis on summative assessment. The latter was a result of the institutional campaign at UH to reduce the workload on staff while enhancing students' learning experience, providing some evidence that the two objectives were not always compatible. However this research does not suggest that summative activities should be avoided, but that the extent of these activities should be carefully engineered and always combined with the formative benefits such as immediate feedback. With regards to the former, staff training is necessary but not the sufficient reason which influences staff competencies with the technology; the role of age, gender, experience and voluntariness of use in this context should be investigated further.

One of the more surprising findings from this study is that despite the overwhelming perception of the ease of use of EVS, the usability could be an issue for students, which implies that training needs to be provided to all students especially if EVS is used in a summative way. Similarly, the institutions should supply accessible handsets to avoid disadvantaging those students with sight and dexterity impairments.

Although students' perceptions on learning benefits and challenges varied across different subject areas, the disparities in views are attributed to the ways in which EVS was used and the experience of students rather than any intrinsic disciplinary differences. Further research in this area could investigate the factors that influence the specific ways in which EVS is used, such as the class sizes, experience of staff, and voluntariness of use. This work provides some evidence that the teachers from non-STEM subjects would benefit from question-writing training and examples of good questions for assessing higher-level knowledge domains in their disciplines.

Wider implications for the large-scale institutional deployment show that the focus for EVS use should remain firmly on the design of pedagogy for enhancing student learning and less on defining reductions in the assessment workload.

References

Bangor, A., Kortum, P., & Miller, J.A. (2008). The System Usability Scale (SUS): An Empirical Evaluation. *International Journal of Human-Computer Interaction*, 24(6), 574-594.

Biasutti, M. (2011). The student experience of a collaborative e-learning university module.

Computers & Education, 57(3), 1865-1875.

Brooke, J. (1986). *System Usability Scale (SUS): A Quick-and-Dirty Method of System Evaluation User Information*. Digital Equipment Co Ltd. Reading, UK.

Bruff, D. (2014). Classroom response system ('clickers') bibliography. Retrieved from the Center for Teaching, Vanderbilt University http://cft.vanderbilt.edu/docs/classroom-response-system-clickers-bibliography/ March 2014.

Caldwell, J. E. (2007). Clickers in the large classroom: Current research and best-practice tips. *Life Science Education*, 6(1), 9-20.

Cardillo, D. (2008). The use of clickers in the law school classroom. *The Law Teacher*, 15(1), 13-14.

Carini, R. M., Kuh, G. D., & Klein, S. P. (2006). Student engagement and student learning: Testing the linkages. *Research in Higher Education*, 47(1), 1-32.

Cohen, L., Manion, L., & Morrison, K. (2000). *Research Methods in Education* (5th ed.) London: Routledge Falmer.

Cutrim, E. S. (2008). Using a voting system in conjunction with interactive whiteboard technology to enhance learning in the English language classroom. *Computers & Education*, 50(1), 338-356.

DeVellis, R. F. (2011). *Scale development: Theory and applications*. Sage Publications.

Draper, S. W., & Brown, M. I. (2004). Increasing interactivity in lectures using an electronic voting system, *Journal of Computer Assisted Learning*, 20(2), 81-94.

Dufresne R.J., Gerace W.J., Leonard W.J., Mestre J.P., & Wenk L. (1996). Classtalk: a classroom communication system for active learning. *Journal of Computing in Higher Education*, 7, 3–47.

Field, A. (2013). *Discovering statistics using SPSS*. (4th ed.) Sage Publications.

Gil-Flores, J., Torres-Gordillo, J. J., & Perera-Rodríguez, V. H. (2012). The role of online reader experience in explaining students' performance in digital reading. *Computers & Education,* 59(2), 653-660.

Hamlen, K.R. (2012). Academic dishonesty and video game play: Is new media use changing conceptions of cheating? *Computers & Education*, 59(4), 1145-1152.

Han, J. H., & Finkelstein, A. (2013). Understanding the effects of professors' pedagogical development with Clicker Assessment and Feedback technologies and the impact on students' engagement and learning in higher education. *Computers & Education*, 65, 64-76.

Han, J. H. (2014a). Closing the Missing Links and Opening the Relationships among the Factors: A Literature Review on the Use of Clicker Technology Using the 3P Model. *Educational Technology & Society*, 17 (4), 150–168.

Han, J. H. (2014b). Unpacking and repacking the factors affecting students' perceptions of the use of classroom communication systems (CCS) technology. *Computers & Education*, 79, 159-176.

Hancock, T. M. (2010). Use of audience response systems for summative assessment in large classes. *Australasian Journal of Educational Technology,* 26(2), 226-237.

Jenkins, A. (2007). Technique and technology: Electronic voting systems in an English literature lecture. *Pedagogy*, 7(3), 526-533.

Kaleta, R., & Joosten, T. (2007). Student response systems: A University of Wisconsin system study of clickers. *EDUCAUSE Research Bulletin*, 2007(10), 1–12.

Kay, R. H., & LeSage, A. (2009). Examining the benefits and challenges of using audience response systems: A review of the literature. *Computers & Education*, 53(3), 819-827.

Lantz, M. E., & Stawiski, A. (2014). Effectiveness of clickers: Effect of feedback and the timing of questions on learning. Computers in Human Behavior, 31, 280-286.

Laxman, K. (2011). A study on the adoption of clickers in higher education. *Australasian Journal of Educational Technology*, 27(8), 1291-1303.

Manisera, M., Van der Kooij, A. J., & Dusseldorp, E. (2010). Identifying the component structure of satisfaction scales by nonlinear principal components analysis. *Quality technology & quantitative management,* 7(2), 97-115.

Mazur, E., & Hilborn, R. C. (1997). Peer instruction: A user's manual. *Physics Today*, 50(4), 68-69.

Micheletto, M. J. (2011). Using audience response systems to encourage student engagement and reflection on ethical orientation and behavior. *Contemporary Issues in Education Research (CIER),* 4(10), 9-18.

Muncy, J., & Eastman, J. (2012). Using classroom response technology to create an active learning environment in marketing classes. *American Journal of Business Education*, (AJBE), 5(2), 213-218.

Simpson, V., and Oliver, M. (2007). Electronic voting systems for lectures then and now: A comparison of research and practice, *Australasian Journal of Educational Technology*, 23(2), 187-208.

Steventon, B., Panesar, S., & Wood, J. (2012). Moving the law school into the twenty-first century–embedding technology into teaching and learning. *Journal of Further and Higher Education*, 38(1), 107-128.

Sun, P. C., Tsai, R. J., Finger, G., Chen, Y. Y., & Yeh, D. (2008). What drives a successful e-Learning? An empirical investigation of the critical factors influencing learner satisfaction. *Computers & Education*, *50*(4), 1183-1202.

Twetten, J., Smith, M. K., Julius, J., & Murphy-Boyer, L. (2007). Successful clicker standardization. *Educause Quarterly*, 30(4), 63-67.

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. B. (2003). User acceptance of information technology: toward a unified view. *MIS Quarterly*, 27(3), 425–478.

Yeh, C. R., & Tao, Y. H. (2013). How Benefits and Challenges of Personal Response System Impact Students' Continuance Intention? A Taiwanese Context. *Educational Technology & Society*, 16 (2), 257–270.

Appendix A. Additional PCA tables

Structure Matrix	Components			
	1	**2**	**3**	**4**
Using the EVS allows problem areas to be identified	**0.835**			
EVS provides an immediate check of students' understanding	**0.832**			0.438
I enjoy using EVS in my learning **	**0.788**	-0.466	0.302	0.724
Viewing responses gives me an idea of how I am doing in relation to others	**0.777**			
I look forward to using EVS in the future **	**0.777**	-0.448	0.310	0.739
I think I would like to use EVS frequently	**0.771**	-0.412		0.669
Using EVS enables me to accomplish learning activities more quickly	**0.758**			0.722
The advantages of using EVS outweigh any disadvantages **	**0.754**	-0.447	0.348	0.606
Responding to questions in class makes me think about the course	**0.746**			
I am very satisfied with class sessions that use EVS **	**0.743**	-0.467	0.341	0.693
The lecturer addressed relevant topics or issues identified by responses	**0.726**			0.418
I found EVS very awkward to use		**0.765**	-0.456	
Sometimes it is not clear what I am supposed to be voting for		**0.734**		
Using EVS can distract from the learning point entirely	-0.566	**0.712**		
Setting up and using the handsets takes too much time in lectures	-0.465	**0.684**		-0.436
I thought that there was too much inconsistency in the EVS functions		**0.678**		
I needed to learn a lot of things before I could get going with EVS		**0.678**		
I think that I would need the support of a technical person to use EVS		**0.631**		
I found EVS unnecessarily complex		**0.62**	-0.527	
Sometimes lecturers seem to be asking questions just for the sake of it		**0.604**		
Often, I do not have enough time to think before I have to vote		**0.535**		
I thought EVS was easy to use			**0.739**	
I felt very confident using EVS	0.441	-0.529	**0.707**	
I would imagine that most people would learn to use EVS very quickly			**0.693**	
The EVS handset is suitable for a student with disabilities			**0.591**	
I found that the various functions in EVS were well integrated	0.533		**0.543**	
I think about the subject outside of the class more in classes that use EVS	0.423			**0.846**
I spend more time discussing the subject in classes that use EVS	0.455			**0.845**
I ask more questions in classes that use EVS				**0.818**
I am motivated to attend classes that use EVS				**0.757**
Using EVS will increase my chances of getting a good grade	0.568			**0.581**

Component Correlation Matrix				
Component	**1**	**2**	**3**	**4**
1	1.000	-.356	.228	.533
2	-.356	1.000	-.254	-.241
3	.228	-.254	1.000	.106
4	.533	-.241	.106	1.000

[i] The full list of bibliographic databases subscribed to by the authors' host institution contains 65 entries.