Lancaster University Module Evaluation System (LUMES)

Best Practice Guide

By Tim Ellis

t.ellis@lancaster.ac.uk

(5)94835

Contents

Brief introduction to the guide

This guide is designed to help departments and departmental staff get the most benefit from the online module evaluation system (LUMES). The guide is separated into sections for ease of use and reference. LUMES is a centrally administered system that is designed to collect and collate student feedback and perceptions of experience regarding a module. It is one of a number of tools that should be used when evaluating a module in its context. Other forms of data include overall performance of students or external examiner reports. Used together all these data sources can be used to fully evaluate a module.

Benefits of online module evaluation

One of the greatest benefits on using the online module evaluation system is that is allows the efficient and systematic collection of feedback from students on all taught modules with savings in administrative time and effort. By automatically processing the setup, email alerts, reminders, collection of response, and report generation, the time saving benefits are significant. As an online system, it also makes the creation, modification and testing of different questions and questionnaires much easier and faster. In addition to this, feedback from departments indicates that the quality of feedback given by students is of a higher quality than that provided when completing a paper based questionnaire. However, the ultimate purpose of the system is to collect student feedback to improve instruction.

How to Configure the Module Evaluation System

In order to configure or modify the module evaluation system, a login is required. This is usually provided as part of the training sessions run by the Learning Technology Group. Sessions are run every term and are widely advertised. To find out when the next sessions will run, contact learningtechnology@lancaster.ac.uk or on the events page of the CELT website (http://www.lancs.ac.uk/celt/celtweb).

This document will not explain in detail how to make all the modifications, but this information can be found here (http://www.lancs.ac.uk/celt/celtweb/staff_module_evaluation).

Factors affecting response rates (and how to improve them)

Response rates are key factor in the success. An online system such as this one, which is completely voluntary and divorced from the face-to-face teaching time, can struggle to match the response rates achieved when evaluation forms are handed out and retrieved in the classroom situation. Nevertheless, data from previous year's online module evaluations have shown that response rates greater than 90% (even up to 100%) are possible using this system. The following factors may influence the response rates of modules:

- **Accuracy of data.** One of the key factors to achieving good response rates is to ensure that the evaluation is performed at the correct time. The timing of the automated module evaluations is based on data in the LUSI (Lancaster University Student Information) database and if this data is inaccurate, then the module evaluation will be initialised at the wrong time (either too late or too early). Students who receive evaluation requests at the wrong time are less likely to complete subsequent evaluations.
- **Ask appropriate questions**. Ensure that the questions that are being asked are appropriate to the module. The core questions that are on every questionnaire are sufficiently generic to match to all modules. Ensuring that the optional questions are matched to the module will increase the chances that the students will answer them in a meaningful way and will not get frustrated that they are asked questions that are not relevant, e.g. do not ask about seminars if the module contained none. The delivery mode specified in LUSI is used to automatically assign an appropriate questionnaire to a module and so ensuring the data recorded in LUSI is accurate will help to deliver the correct questionnaire to the correct type of module (e.g. dissertation based modules or projects).
- **And not too many.** Another factor achieving good response rates is not to ask too many questions. A respondent who is faced with an extended or complex set of questions is more likely to abandon the questionnaire. The default questionnaire contains approximately 10 questions and it is advised that unless absolutely necessary, the number of questions should be restricted to 10 or fewer.
- **Departmental/teaching staff engagement/encouragement.** A major factor in achieving good response rates is to engage with the system and discuss it with your students. As the email requests for module evaluation come from a centralised service, the students will be less inclined to respond if they think it has nothing to do with their department. By pre-warning students regarding the requests and encouraging them to complete the evaluations, there is a higher chance that the students will comply with the request. Posting information on noticeboards, in handbooks and on LUVLE pages, the students will see module evaluation as a departmental process and therefore be more likely to complete them.
- **Show how the data is used.** Linked to the previous point, students will be more likely to complete module evaluations if they can see that their input is used and not simply filed away. Posting numerical results and/or student comments from module evaluations along with departmental responses for students to view, or discussing the results in staff-student committees will re-affirm the importance that department staff place on student feedback. A very powerful comment is to simply reflect on an example from the previous survey that caused a change in the way the current module was taught. If you know when and how data from LUMES will be made available to students by you, your department or your faculty, tell them.

- **Give enough time (but not too much).** By default respondents are given 2 weeks to complete an evaluation which includes regular reminders (to a maximum of 3) if they have not completed. This is longer for modules at the end of michaelmas and lent terms to allow students leaving early a chance to reply when they return at the start of the next term. Evidence gathered so far shows that roughly $^{2}/_{3}$ of the final response rate will be achieved in the first week. Every subsequent week of evaluation brings diminishing returns with two weeks considered to be optimal. Choosing to reduce the amount of time students are given to complete an evaluation may reduce the final response rate for the evaluation.

Getting the most out of the module evaluation reports

Before an evaluation begins, it is important for the department or course convenor to decide upon what questions the module evaluation will contain. Questions should be focused on what, if anything, needs to be changed to improve the module. Knowing what data is wanted out of the system will then drive what questions are added (*or removed*) from the questionnaire. These questions may arise from issues identified through the Annual Teaching Review or Periodic Review processes at a departmental level, or may be more ad-hoc "would like to know what the students thought of..." question at an individual module level (for instance: to test student response to a recent change or innovation in the teaching or assessment approaches) . Either way, it is important to have an idea of what information is required as a result of each questionnaire before the process begins. For more information about how to amend questionnaires, see the section on "How To Configure the Module Evaluation System".

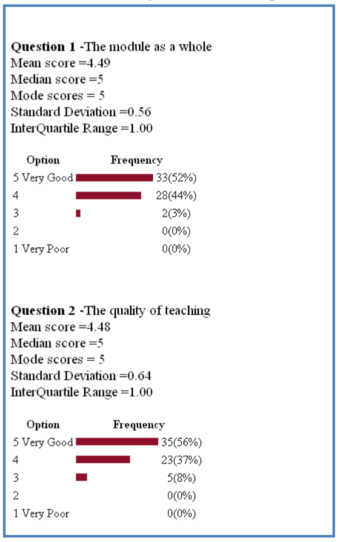
The second question that requires an answer is what information is required on the report and how it should be presented. There are three main areas that affect the way that the data is presented in a report in more than a simple aesthetic manner.

- The addition of statistical data regarding each question. (The interpretation of statistical data can be difficult and a more in depth analysis of the statistical data is presented in the next section).
- The addition of demographic breakdowns. In relation to module evaluation, demographics refer to any event specified in LUSI which is not a lecture. For example, any tutorial or seminar group to which a student belongs will be passed to the results as a demographic. This gives the option to break down the results by each separate tutorial or seminar group, which may help to analyse differences between sections of the module or responses to different tutor contributions. This can be presented alongside the overall question data or a separate page for each demographic group can be created which can then be distributed to the individual tutor or group leader.
- The option to generate a detailed breakdown of the results. This option will create a separate page for each evaluation submission showing how each individual respondent completed the questionnaire. This breakdown can be used to cross reference particular comments or scores between questions and can be useful in spotting aberrations or anomalous results. If more complex cross referencing is required (or imports to third party tools such as SPSS) then an Export to Excel option can be requested to the Learning Technology Group (learningtechnology@lancaster.ac.uk).

Interpreting module evaluation statistics

In the module evaluation reports, it is possible to include statistical analysis of each Likert scale question (http://en.wikipedia.org/wiki/Likert_scale). These statistical parameters include: mean, median, mode, standard deviation and inter-quartile range.

Extract from report showing statistical data



Question 1 -The module as a whole
Mean score =4.49
Median score =5
Mode scores = 5
Standard Deviation =0.56
InterQuartile Range =1.00

Option	Frequency
5 Very Good	33(52%)
4	28(44%)
3	2(3%)
2	0(0%)
1 Very Poor	0(0%)

Question 2 -The quality of teaching
Mean score =4.48
Median score =5
Mode scores = 5
Standard Deviation =0.64
InterQuartile Range =1.00

Option	Frequency
5 Very Good	35(56%)
4	23(37%)
3	5(8%)
2	0(0%)
1 Very Poor	0(0%)

Mean and Standard deviation

Mean and standard deviation are two of the most recognisable statistical measurements and most people are familiar with their meaning. The mean specifies the average score (or value) or a measurement, and the standard deviation refers to the average spread of the data from the mean. Mean and standard deviation are included in the standard report by default as a general measure of student responses to a question but it should be noted that caution should be used when quoting these values as meaningful statistics.

The problem in using mean and standard deviation on Likert scale questions is due to the nature of the data in question. The data obtained from a Likert scale question is an example of *ordinal* data where each data point (choice on the scale) is a defined category (Strongly Agree, Agree, etc) and they are generally arranged in an order. A score may be assigned to each category (5 = Strongly Agree, 4 = Agree, etc) which is then used to calculate the mean and standard deviation. Experts in the field point out that that using these statistics on ordinal data such as Likert scale questions is not valid because each point on the scale is not necessarily equally distant from the one that preceeded it and the one that followed it. For example the distance between Strongly Agree and Agree may not be the same as the distance between Agree and Neither Agree Nor Disagree. To address this problem, the module evaluation system does not label each choice but only gives the top and bottom of the scale (Very Good & Very Poor; Strongly Agree & Strongly Disagree) thereby leaving the respondent to interpret the points in between to be equally separated. This gives the statistical data increased validity but caution should still be applied.

Mode, Median and Inter-quartile range

These statistical parameters are less well known but lend themselves better to Likert scale data as they do not involve any averaging calculations. Mode is simply the choice which received the most selections (most popular). Median measures the "middle" value. If there were 9 results and all the selections were ranked from lowest to highest then the median would be the 5th result. Inter-quartile range is a measure of the spread of data and gives a value which constitutes the range of the middle 50% of the data. A low inter-quartile range indicates data which is tightly grouped.

Other useful parameters

Another useful parameter often used when presenting the results of Likert scale questions is to quote the percentage response of each option. This can be done using individual choices (75% chose option 5, 15% chose option 4, etc) or the positive choices (4&5) and the negative responses (1&2) can be grouped so that a grouped response can be quoted; "90% of responses were either Good or Very Good".

Interpretation and usage

Each of these statistical parameters can be used in a number of ways to measure the quality of a module and how they are used is entirely the responsibility of the staff member or department. To improve the quality of the data presented, any statistics used should be consistent whenever used in a comparative. *It should also be noted that the quality of any statistics is entirely dependent on the size of the population (i.e. the number of responses) with a higher number of responses giving better quality (and reliable) results than a lower one.* Any report containing fewer than 10 results contains a disclaimer that the statistics are less reliable. There is debate over how many results constitutes a "minimum" but anything over 25 results can be considered a good number of results. This is why it is important for staff to encourage 100% participation from their students.

Comparisons between modules should be done cautiously as many parameters can affect how a module is rated by students (e.g. difficult modules may be rated lower than easy ones, compulsory modules may be rated lower than optional modules).

Further analysis of the data is possible and it is possible to test whether the difference between module-module scores or module year-year scores is significantly different. A number of statistical tests can be applied to the data as described here (http://en.wikipedia.org/wiki/Likert_scale).

What to do in the event of receiving abuse, harassment or un-warranted criticism

Students are warned at the beginning of the process about not writing comments that could be considered abusive or harassing and so the instances of such comments being made are very rare. Regardless of the rarity, the university takes such things seriously and therefore, any member of staff who finds a comment in a report generated by LUMES which they consider to be abusive, harassing or un-warranted criticism shall consult their Head of Department (HoD). If the HoD concurs with the assessment of the comment, they will contact the Learning Technology Group (LTG). LTG will not make any subjective decisions regarding the nature of comments or whether they constitute abuse, harassment or un-warranted criticism.

The HoD will forward a formal request either in writing or via email to learningtechnology@lancaster.ac.uk with the following information:

1. The module mnemonic, and
2. Either
    a. A copy of the report in question with the comment highlighted, or
    b. A word for word reproduction of the comment

This will allow LTG to find the comment in the database quickly and effectively.

On receiving the request, LTG will:

1. Disassociate the complete set of responses submitted by the respondent responsible for the comment (not just the comment) from the data set used in the evaluation.
2. Generate a new report for the module which will be sent to the individuals specified in LUMES as recipients for the report for this module. The report will explain that this new report should replace the original, and the original report (and any copies) must be destroyed.
3. The original set of responses will not be initially deleted from the database, but will be disassociated from the module in question and a record kept of the transaction. This is in case further action is required or an appeal is lodged by other parties. Only select members of LTG will have knowledge of retrieval of the disassociated responses and they will only be retrieved on the request from senior officers of the university.

Conclusion

In conclusion, the Lancaster University Module Evaluation System, LUMES, is designed to help you get the feedback you need to improve instruction. The tedium has been removed to allow you to focus on three things: asking the questions you need to make improvements, engaging as many students as possible to increase the reliability of the data produced, and collegially arriving at curricular changes based on student feedback.