Using instant student satisfaction ratings to investigate large scale practical laboratory teaching

*Adam Funnell
Multidisciplinary Engineering Education
University of Sheffield
Sheffield, United Kingdom
*a.funnell@sheffield.ac.uk https://orcid.org/0000-0002-1119-3528

***Abstract*—This Research Work in Progress paper describes the analysis of a rich dataset of student satisfaction with individual lab teaching sessions. In the department of Multidisciplinary Engineering Education, University of Sheffield, a team of dedicated staff design and deliver a range of practical activities across engineering disciplines. Each laboratory activity is meticulously planned and designed, but the broader environmental factors of large scale practical teaching (e.g. class size, timing etc.) are generally fixed, constrained or not considered. Broad module surveys can capture free-form feedback, but this data lacks the granularity to explore individual lab sessions.**

To measure instant student satisfaction with every individual laboratory activity, data was gathered using tablet computers situated at the exits to the laboratory. Students were invited to rate their satisfaction on leaving the session using a 4-point "smiley-face" scale. Over 6 weeks, 95 individual lab sessions were captured, covering three different teaching environments: a 144 seat electronics lab, a 15 seat motors and machines lab, and a teaching cleanroom. The simplicity of the response method elicited a reasonable overall response rate of 43%.

Although the 4 point rating scale does not permit in-depth analysis of the reasons for each student's feedback, whether positive or negative, it does allow a search for correlation between a range of environmental factors. The effect of the number of students per session, and number of students per postgraduate teaching assistant on student satisfaction is considered. The effects of time of day, day of week and duration of session are also explored, alongside the timekeeping of the session e.g. did the laboratory activity take longer or shorter than the timetabled duration. Due to the multidisciplinary engineering teaching approach, the same lab session is often delivered to several different discipline cohorts and years of study - their relative satisfaction with the same practical activity is compared.

***Index Terms*—laboratory teaching, student feedback**

I. INTRODUCTION

Well designed student evaluation of teaching can provide valuable feedback on student perceptions of their learning [1], but care must be taken to ensure that the data gathered is a valid representation of the student experience, and that the data is not used inappropriately for institutional decisions on staffing [2]. A wide range of methods are available to evaluate teaching, from paper based questionnaires, through online surveys, to student focus groups [3], [4]. Institutional level approaches to gathering feedback can create effective frameworks to ensure the process is well resourced and supported, but the size of questionnaire and lack of direct connection to actions taken can create barriers to student responses [5].

Feedback is generally collected in surveys at module level, although this does reflect the multidimensionality of teaching styles within a module, such as lectures, tutorials and laboratory sessions [6]. In addition, student evaluation surveys can conflate students' prior expectations of their course with teaching quality, institutional satisfaction, and overall student experience [7]. Student satisfaction has been shown to be affected by all of their individual interactions with the institution [8]. These theories suggest that frequent surveys may be the most effective method of obtaining student feedback [9].

For an effective and efficient multidisciplinary education environment, laboratory teaching for undergraduate engineers at the University of Sheffield is provided in a dedicated building by specialist practical teaching staff [10]. Engineering teaching laboratories are a completely different teaching environment to lecture theatres or classrooms, and as environment has a strong impact on student evaluation of teaching [11], laboratory sessions deserve independent evaluations.

By measuring students' satisfaction with each laboratory teaching session as they leave the room, feedback can be directly collated with the exact teaching session it relates to. Single-question exit surveys may not allow time for carefully considered responses, but even end-of-module surveys are often rushed and dominated by emotional thoughts [12]. An investigation into frequent student feedback found few differences in students' ratings when comparing frequent surveys mid-semester to single surveys at the end of a course [13], although this work did not separate out different teaching environments within a course (e.g. labs vs. lectures).

Initial analysis is presented here of student feedback data from laboratory sessions during the first six weeks of the autumn semester in 2019. This work attempts to correlate the student feedback ratings with various environmental factors, to see if meaningful general conclusions can be drawn about the provision of large scale laboratory teaching.

II. DATA CAPTURE METHODS AND LIMITATIONS

To capture data, tablet computers were placed on stands at the exits to laboratory spaces, and set up to gather student feedback data on each individually timetabled teaching session. Four options were presented for students to press on the tablet screen, in the form of smiley faces, as shown in Fig. 1.

Fig. 1. The four possible smiley face options shown to students

Students can press a single face to show their feelings about the session, before the system screen freezes for one second with a ”Thank you” message to reduce the likelihood of multiple button presses per student. No further guidance is given on-screen to the students on how they should record their feelings. Students were reminded verbally during the lab activities to press a button upon leaving the room to give their feedback, but were not forced to do so or monitored while they exited. The data capture was completely anonymous; the tablet captured each button press with a precise timestamp, but stored no further data on the student.

For data analysis, numerical values are assigned to each face, of 4 for the happiest, down to 1 for the least happy. In the subsequent discussions, the data is either analysed in terms of ”session” ratings (aggregating all button presses logged as part of a specific lab teaching session); or ”student” ratings where each individual button press is assigned into a wider category e.g. all ratings received on Monday mornings.

Limitations of this data include the lack of separation of feedback from different sessions happening in the same laboratory simultaneously (any student can leave feedback on any tablet, not just the tablet assigned to their activity), and the ease with which a student could leave multiple responses (by simply waiting for the ”thank you” message to disappear from the screen). This dataset should therefore not be considered high quality for any individual session, but these issues are systematic and affect all sessions equally.

The limits of this method of data capture are clear. Capturing satisfaction feedback is efficient and allows staff to review frequent feedback instantly; students have positive perceptions of teaching staff who operate mid-semester reviews of teaching effectiveness [14]. However, the data cannot be used for measuring true teaching quality, and is not a measure of staff performance; although the data can reveal students’ general satisfaction with each lab session, it tells us nothing at all about why the students felt that way. Specific reasons for individual sessions scoring relatively high or low satisfaction scores, as measured by this method, can only be fully analysed with contextual observations from staff or students present at the time. This context could be provided by further free-text surveys, however research into sentiment analysis from the large textual datasets of module surveys generally produces general themes rather than specific insight [15], [16].

This paper therefore only attempts to correlate bulk data with environmental factors to explore general trends; the potential use of this data as a feedback tool to improve individual lab sessions is left to future work.

This paper considers a dataset from three laboratory spaces:

- A general electronics and control laboratory, seating 144 students with standard electronics workbench equipment and mechanical control systems for analysis;
- A motors and machines laboratory, seating 15 students with industrial machines and drives test benches; and
- A teaching cleanroom, hosting 16 students, with a range of specialist workstations for micro-scale fabrication

III. Correlation With Time

The mean score for all sessions taught over the six week period was 3.58 out of 4, with a standard deviation of 0.81. This score shows that the students are broadly satisfied with their laboratory experiences. The mean rating for each session is plotted over time in Fig. 2. There is no discernible trend to the mean session rating over the six week period, with least squares regression producing a completely horizontal trendline across the whole time series (not shown in Fig. 2).

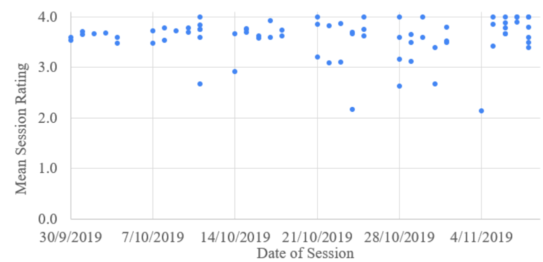

Fig. 2. The mean satisfaction rating per session over a six week period

To determine if students prefer any particular time of week for their lab sessions, all individual student responses were aggregated, then separated into 10 timeslots to represent the session time in the week (i.e. Monday morning, Monday afternoon etc.). The mean of all student ratings in each timeslot is shown in Fig. 3, with standard deviation error bars.

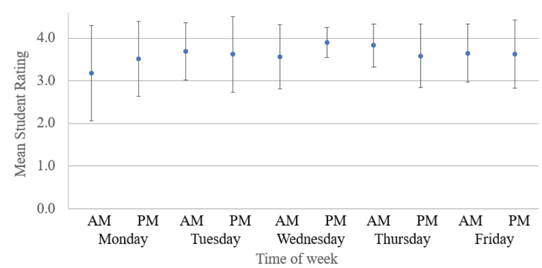

Fig. 3. The mean satisfaction rating per student for each time of day

The large standard deviations (35% in the worst case) show that the results are not statistically significant, although some general hypotheses can be made to explain the subtle differences. The highest mean score was for Wednesday afternoons, which are generally optional sessions for interested and committed students to work on personal projects or for their own personal development, and are likely to result in high self-satisfaction. Monday mornings received the lowest mean score, but that timeslot is often unpopular with students following their weekend activities [17], and early morning lectures have previously been shown to receive lower student feedback scores [18]. There are few actions that can be taken to reduce Monday morning teaching; for high utilisation of laboratory spaces, all available timeslots must be used.

IV. Correlation With Duration

Lab sessions can take a variety of durations, from 1.5 hours up to a maximum of 4 hours. This allows flexibility to provide short lab sessions illustrating single points from lectures, alongside longer free-form design-and-build exercises. The mean score per session is plotted against the timetabled session duration in Fig. 4. More sessions are 3 hours in duration in this dataset than any other length, so it is difficult to draw any conclusions on correlation with session duration.

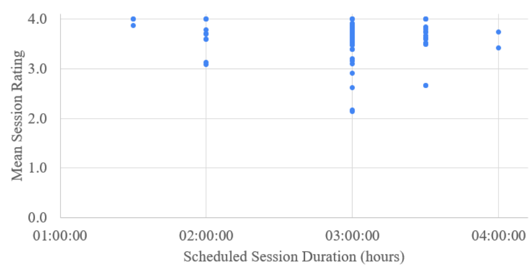

Fig. 4. The mean satisfaction rating per session, sorted by session duration

In practice, sessions may run longer or shorter than their scheduled duration, and this may affect the student experience. Satisfaction may either increase or decrease with time taken, for example, students may be unhappy at needing to spend extra time to finish a task, or they may be enjoying a task and interested to perform extra work. To capture the variation in actual lab duration compared to the scheduled timetable, the timestamp of the last student rating for each session is taken to be the actual time that the session finished. The mean session rating is plotted against the ratio of last rating time to the scheduled lab duration in Fig. 5 i.e. if the last student rating was left at exactly the scheduled finish time, the ratio is 1.0, while if the last student rating timestamp was 30 minutes later than the scheduled finish time of a 2 hour lab, the ratio is 1.25.

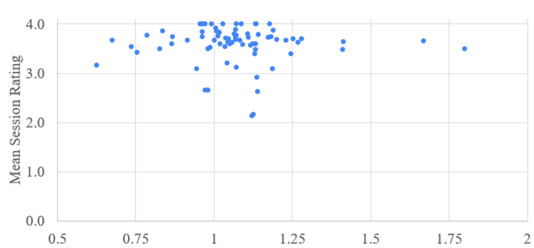

Fig. 5. The mean satisfaction rating per session for ratio of scheduled session duration to actual duration as measured by last feedback timestamp

This measure of actual lab finish time may not be accurate since some students may finish late but decline to leave feedback, but it is a reasonable lower bound estimate for the actual session duration. In addition, this analysis does not account for the speed at which students finish - in some sessions all students finish within minutes of each other, while in other sessions the first students may complete the tasks an hour before the last. Two data points have been removed from this analysis where the finish time was greater than 2x the original scheduled time - this is likely to be an error in operating the data capture tablets.

The highest mean session ratings are only found for ratios between 0.95 and 1.2, which suggests that students are most satisfied when lab sessions run approximately to their original schedule. There is a fall in the mean session rating for lab sessions that are either substantially longer or shorter than scheduled, supporting the hypothesis that students like sessions which match their expectations for time commitment.

V. Correlation With Student and Staff Numbers

To produce multidisciplinary teaching at scale, laboratory sizes must be large, allowing entire cohorts of students to participate in laboratory activities simultaneously for close alignment between practical work and lectures, workshops and seminars [19]. Although the maximum capacity of the largest lab studied in this work was 144 seats, the space was often configured for multiple smaller classes to take place simultaneously. In the two smaller labs, the student capacity was far smaller due to limited sets of specialist equipment. The mean session rating is plotted against student numbers per session in Fig. 6. There is a very slight trend for student satisfaction to decrease as session size increases.

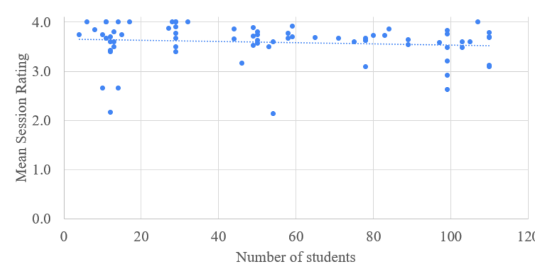

Fig. 6. The mean satisfaction rating per session for number of students in the session

It is only possible to teach laboratory classes with the support of a large number of graduate teaching assistants (GTAs), who are often the first line of contact between students and staff in the laboratory. The ratio of GTAs to students in the lab is determined by the availability of trained staff and the complexity of the experiment among many other factors. The mean session rating is plotted against the number of students per GTA in each session in Fig. 7. Discounting the single data point at 75 students per GTA, which would vastly skew the trendline, there is a negative trend in mean session rating as the number of students per GTA increases, although the trendline is plotted with low confidence ($R^2 = 0.04$). This trend is likely due to the increased waiting times for help if many students require assistance from only a few GTAs.

VI. Correlation With Discipline Cohort

Within the first few weeks of the semester, students from many different engineering disciplines participate in a ”work-station familiarisation” lab session, to introduce them to

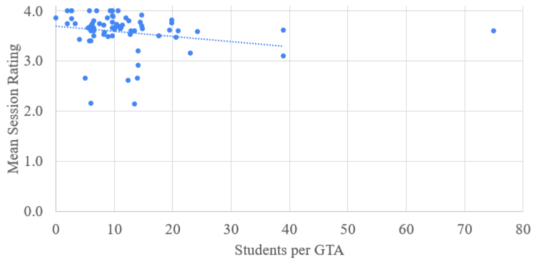

Fig. 7. The mean satisfaction rating per session for number of students per Graduate Teaching Assistant

standard workbench equipment and safe working practices. All of the students taking part in this activity are first year students, with the exception of mechanical engineering second year students. The activity is identical for all disciplines, allowing a search for correlation between programme of study and satisfaction with the session. The mean student rating (aggregated across multiple sessions where there are more students in the cohort than could be accommodated in a single session) for each discipline is shown in Fig. 8.

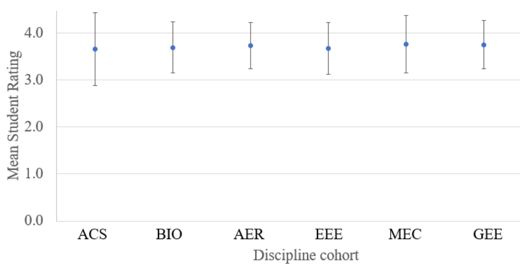

Fig. 8. The mean satisfaction rating per session split by discipline cohort. ACS = Automatic Control and Systems; BIO = Bioengineering; AER = Aerospace; EEE = Electrical and Electronic; MEC = Mechanical; GEE = General.

The differences between the disciplines are subtle; mean scores range from 3.65 to 3.75. Although for this introductory lab activity the scores are similar, this method of analysis could be a useful tool to analyse if other lab activities later in the students' programmes are more satisfying than others depending on the discipline context they are set in.

VII. Correlation With Weather

This final analysis looks for correlations between external factors and student satisfaction data, using the fine granularity time resolution of this particular dataset. Prior work has explored how some methods of student evaluation actually measure the "realized utility" of the courses, rather than the quality of the teaching itself [20]. External factors such as weather are known to affect perception of utility, and in analysis of a small dataset in [20], a positive correlation between weather and student evaluation scores is argued to show that student ratings depend more on the perceived utility of the course rather than "teaching quality". Given the fine granularity of the temporal resolution of this dataset, this argument can be explored further. Weather data is taken from a nearby weather station; the mean session rating is plotted against both the total rainfall and the average temperature in the 24 hour period around the lab session in Fig. 9.

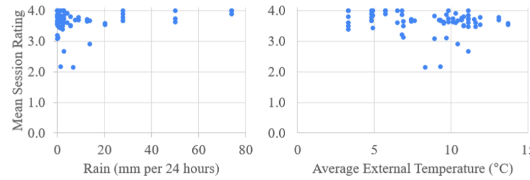

Fig. 9. The mean satisfaction rating per session for external weather conditions: left, rainfall and right, temperature

This data shows no correlation between the external weather and student satisfaction ratings. This data does not support the premise from [20], that evaluation scores are dominated by session utility. However, this evidence alone is not sufficient to claim this method is a valid measure of teaching quality. Weather is just one example of external influences that may affect student satisfaction; public transport, campus environment and activity levels, even catering data could all be correlated with the rich dataset from this system in future work.

VIII. Discussion and Conclusions

By using tablet computers at the exits from laboratories, instant feedback was gathered on student satisfaction. The data gathered has severe limits in analysing individual sessions but could show general trends over time.

The actions to take forward to continuously improve student experience have three parts. Firstly, the only distinct conclusions from this dataset are that students appreciate laboratory sessions which closely fit their timetabled duration, and smaller class sizes and staff/student ratios tend to result in more satisfied students. Staff should design and run laboratory sessions which meet these aims.

In addition, the attempts at correlation presented here do not particularly support the deconstruction of student satisfaction into any of the categories studied. Prior work has also attempted an empirical breakdown of reasons for overall satisfaction [21], although that work used a thorough questionnaire rather than data sources, with only weak correlations observed. The validity of this approach should be verified by a secondary study employing further surveys (such as questionnaires and focus groups), although it may be challenging to carry out these methods as frequently and time-aligned to the teaching session as the original capture method.

Thirdly, to try and reduce bias within the results, the data capture method should be expanded to allow separation of key characteristics shown to impact on student evaluations of teaching, such as gender. At present the responses are acquired with complete anonymity, due to the rapid nature of a single button press to provide feedback. It could be possible to use multiple tablets, and ask students to only respond on corresponding tablets to separate key characteristics. Alternatively the survey design could be adapted to ask multiple questions and collect further contextual data, at the risk of reduced response rates and incomplete data gathering.

REFERENCES

[1] F. Hammonds, G. J. Mariano, G. Ammons and S. Chambers, "Student evaluations of teaching: improving teaching quality in higher education", Perspectives: Policy and Practice in Higher Education, vol. 21 no. 1, 2017, pp. 26-33. doi: 10.1080/13603108.2016.1227388

[2] J. Esarey and N. Valdes, "Unbiased, reliable, and valid student evaluations can still be unfair", Assessment & Evaluation in Higher Education, 2020 pp.1-15. doi: 10.1080/02602938.2020.1724875

[3] N. K. Mandal, "Importance of student feedback in improving mechanical engineering courses", International Journal of Mechanical Engineering Education, vol. 47, no. 3, 2019, pp. 227–245. doi: 10.1177/0306419018759800

[4] J. L. Chang, M. J. Piket-May and J. P. Avery, "Using active student feedback in the learning environment", FIE '98. 28th Annual Frontiers in Education Conference. Moving from 'Teacher-Centered' to 'Learner-Centered' Education. Conference Proceedings (Cat. No.98CH36214), Tempe, AZ, USA, vol. 2, 1998, pp. 643-646. doi: 10.1109/FIE.1998.738760

[5] C. S. Nair, A. Patil and P. Mertova, "Enhancing the quality of engineering education by utilising student feedback", European Journal of Engineering Education, vol. 36, no. 1, 2011, pp. 3-12. doi: 10.1080/03043797.2010.518231

[6] H. W. Marsh, "Students' evaluations of University teaching: Research findings, methodological issues, and directions for future research", International Journal of Educational Research vol. 11, no. 3, 1987, pp. 253-388. doi: 10.1016/0883-0355(87)90001-2

[7] S. L. Appleton-Knapp and K. A. Krentler, "Measuring student expectations and their effects on satisfaction: The importance of managing student expectations", Journal of marketing education, vol. 28, no. 3, 2006, pp.254-264. doi: 10.1177/0273475306293359

[8] S. Wilkins and M. S. Balakrishnan, "Assessing student satisfaction in transnational higher education", The International Journal of Educational Management, vol. 27, no. 2, 2013, pp.143-156. doi: 10.1108/09513541311297568

[9] T. M. Winchester and M. Winchester, "Exploring the impact of faculty reflection on weekly student evaluations of teaching", International Journal for Academic Development, vol. 16, no. 2, 2011 pp. 119-131. doi: 10.1080/1360144X.2011.568679

[10] P. Kapranos, "The Interdisciplinary Future of Engineering Education: Breaking Through Boundaries in Teaching and Learning", Routledge, 2018. doi: 10.4324/9781351060790

[11] M. C. Hill and K. K. Epps, "The impact of physical classroom environment on student satisfaction and student evaluation of teaching in the university environment", Academy of Educational Leadership Journal, vol. 14, no. 4, 2010, p65.

[12] M. B. Pinto and P. Mansfield, "Thought Processes College Students Use When Evaluating Faculty: A Qualitative Study", American Journal of Business Education (AJBE), vol. 3, 2010. doi: 10.19030/ajbe.v3i3.399.

[13] K. S. McOwen, J. R. Kogan and J. A. Shea, "Elapsed Time Between Teaching and Evaluation: Does It Matter?", "Academic Medicine", vol. 83, iss. 10, 2008, pp. S29-S32. doi: 10.1097/ACM.0b013e318183e37c

[14] M. J. Brown, "Student Perceptions of Teaching Evaluations", Journal of Instructional Psychology, vol. 35, no. 2, 2008, pp. 177-181.

[15] S. Rani and P. Kumar, "A sentiment analysis system to improve teaching and learning.", Computer, vol. 50, no. 5, 2017. pp. 36-43. doi: 10.1109/MC.2017.133

[16] M. A. Ullah, "Sentiment analysis of students feedback: A study towards optimal tools", 2016 International Workshop on Computational Intelligence (IWCI), IEEE, 2016, pp. 175-180. doi: 10.1109/IWCI.2016.7860361

[17] C. Paisey and N. J. Paisey, "Student attendance in an accounting module – reasons for non-attendance and the effect on academic performance at a Scottish University", Accounting Education, vol. 13, no. sup1, 2004, pp. 39-53. doi: 10.1080/0963928042000310788

[18] R. G. Tobin, "Too Early for Physics? Effect of Class Meeting Time on Student Evaluations of Teaching in Introductory Physics", The Physics Teacher vol. 55, 2017 p276. doi: 10.1119/1.4981033

[19] M. Judge, "Large-scale laboratory teaching for 1st year EEE undergraduates", International Journal of Electrical Engineering Education, vol. 54, no. 2, 2017, pp. 164-177. doi: 10.1177/0020720916688487

[20] M. Braga, M. Paccagnella and M. Pellizzari, "Evaluating students' evaluations of professors", Economics of Education Review, vol 41, 2014, pp.71-88. doi: 10.1016/j.econedurev.2014.04.002

[21] J. Wiers-Jenssen, B. R. Stensaker and J. B. Grøgaard, "Student satisfaction: Towards an empirical deconstruction of the concept", Quality in higher education, vol. 8, no. 2, 2002, pp.183-195. doi: 10.1080/1353832022000004377