What factors influence student satisfaction with module quality?

A comparative analysis of the determinants of module satisfaction in undergraduate and taught postgraduate business school students

Dylan Sutherland, Durham University

Philip Warwick, Durham University

John Anderson, University of Northern Iowa

Abstract: Understanding the links between student satisfaction and the quality of teaching is becoming of ever greater concern to UK business schools. Here we explore the determinants of overall student module satisfaction, as well as their relative importance, by comparing the module evaluation questionnaire (MEQ) responses of undergraduate (UG) and postgraduate (PG) students in a UK business school. Based upon econometric modelling of the MEQ averages from 470 UG modules (21,096 student responses) and 93 masters modules (4,429 responses) our results identify several key factors driving overall satisfaction in both groups. The helpfulness of lectures and seminars, involving direct student-teacher contact time, is the most important determinant of satisfaction. This is followed, in order of importance, by: the degree to which a module integrates well with other elements of the course; the usefulness of supporting on-line materials; and the appropriateness of summative assignments. Useful readings and feedback on formative assignments appear weaker drivers of overall satisfaction, particularly for PG students, as is the negative impact of perceived module difficulty. In general, however, the drivers of satisfaction in PG and UG groups are similar. We discuss the implications of the findings for academic staff, business school managers and the HE sector.

Key Words: Module evaluation, MEQ, student satisfaction, business school, regression analysis

INTRODUCTION

As domestic and international competition in higher education intensifies there is increasing pressure on universities to increase student satisfaction with teaching, as it is reflected in improved performance in module evaluation questionnaires (MEQs) and university rankings. University managers wish to see improvements in student satisfaction so they can boost rankings and attract more, better qualified students. Teaching staff, aware their career prospects are increasingly determined by MEQs, are keen to understand what influences MEQs and how to improve them. Government policy-makers, who oversee university funding, are also in future more likely to reward universities with better teaching performances by allocating them larger shares of public funds (or allowing them to raise student fees). They too, in a bid to create viable teaching evaluation frameworks, such as the UK's Teaching Enhancement Framework (TEF), wish to understand what factors shape student satisfaction with taught modules. Student satisfaction with teaching has recently been found to be among the most important determinants of programme 'value for money' in UK universities (Neves and Hillman 2016). The UK higher education sector, recognising this, and increasingly being driven by powerful market forces, goes to considerable efforts to improve student satisfaction of teaching. Developing a better understanding of the determinants of satisfaction could therefore be useful for university and departmental managers, module leaders, lecturers and university funding agencies (Ward & Shortt 2012). In this light, it is logical to analyse MEQ data so as to try to better understand what drives student satisfaction and also whether there are differences in the types and magnitudes of drivers between different types of students, such as undergraduate (UG) and postgraduate (PG) groups. By doing so, we can gain better insights into how student satisfaction can be most efficiently improved, as well as the potential dangers and perhaps unforeseen consequences of relying too greatly upon module evaluations. The purpose of this paper is therefore to empirically explore the determinants of perceived module quality using around 25,000 undergraduate and postgraduate student responses from an internationally accredited business school.

We first develop hypotheses regarding the relative importance of determinants of satisfaction with module quality for both PG and undergraduate UG students. We then explain the methodology. The results and discussion follow. We find close similarities in the drivers of satisfaction for both PG (taught Masters level) and UG students. Direct contact time with teaching staff has the greatest impacts on overall satisfaction for both

groups. Some indirect factors, however, also play important roles for both groups, including module integration and usefulness of on-line resources. Surprisingly, however, we find useful reading and formative feedback to be comparatively unimportant drivers of overall student satisfaction at both UG and PG level. Perhaps of most interest (and some concern), we find a strong negative relationship between module difficulty and perceived satisfaction, raising the question of whether excessive dependence on MEQs to evaluate academic staff may ultimately lead to the dumbing down of the UK higher education sector.

LITERATURE REVIEW AND HYPOTHESIS DEVELOPMENT

What determines student satisfaction with taught modules in business school higher education modules? Empirically focused research undertaken in the general area of student satisfaction now dates back over three decades. It still, however, remains a relatively narrow field within education related research. Most studies, like ours, have drawn from student evaluation data (Broder & Dorfman, 1994; Hearn, 1985; Krahn & Bowlby, 1997; Nadiri, Kandampully, & Hussain, 2009; Neumann & Neumann, 1981; Rienties, Li, & Marsh, 2015). They have, however, been undertaken at a number of different levels of analysis: some consider evaluations of entire courses, programmes or the university experience (Filak & Sheldon, 2003; Rienties et al., 2015); some module satisfaction (Broder & Dorfman, 1994; Rienties et al., 2015); and others, filling different niches, have looked at determinants of curriculum satisfaction (Tessema et al, 2012). Research on satisfaction determinants in specific subjects or fields also exists. This includes psychology (Green, et al 2015), sports (Poppet al, 2015), music (Serenko, 2011) and also in business school related courses. Indeed, we identified seven business school related studies, making it the most studied subject area (Bennett, 2003; DeShields et al, 2005; Douglas et al 2014; Hill et al, 2003; Letcher & Neves, 2010; Malik, Danish, & Usman, 2010; Shurden et al, 2016).The focus of most of these satisfaction studies is at the undergraduate level, involving US and UK based students (Bennett, 2003; Douglas et al, 2014) although other countries have been studied, for example Greece (Nadiri et al, 2009), Pakistan (Malik et al, 2010) and the UAE (Dodeen, 2016).

At a conceptual level, two prominent areas have attracted interest within educational research. Firstly, what are the most important determinants of student satisfaction? Or, as Hearn (1985) puts it in his pioneering study: "how do students weight the various domains of satisfaction and dissatisfaction (e.g. faculty availability, faculty

teaching ability) in arriving at their levels of overall program satisfaction?"(Hearn,1985: 415).   Understanding the key satisfaction drivers, as mentioned, could be of useful commercial value to higher education institutions. It could help them to better attract and retain students. From a pedagogical perspective, moreover, it is also very important, as satisfaction is closely related to motivations for studying. This, in turn, positively influences learning outcomes (Hearn, 1985).

Secondly, once we have established what the general drivers of satisfaction are, how do the weights on these drivers vary across different groups, such as student groups in different subjects or disciplines or at different levels (i.e. PG and UG)? The processes of satisfaction formation, for example, have been found to be different across different subject groups (Hearn, 1985; Neumann & Neumann, 1981). Indeed, recent work by Neves and Hillman (2016 and 2017) suggests that satisfaction continues to show marked variation between subject areas in the UK. Assuming different student groups are satisfied in similar ways, therefore, oversimplifies matters. The different weightings found on satisfaction drivers between subject groups have been referred to as to as "field differences" by Hearn (1985). The "field' here refers to different subject fields, or academic discipline generally. Hearn's (1985) early empirical analysis, for example, compared satisfaction drivers in six different fields. The idea (and associated methodologies), however, can be equally applied to other types of groups – such as UG and PG students, as we do here. Since Hearn (1985), there has been comparatively little research on field differences or, more generally, how the drivers of student satisfaction vary across different groups (Broder & Dorfman, 1994).[1]

In summary, in the same vein as the student satisfaction literature, we look to establish both the significance and magnitude of the drivers of student module satisfaction. We do so for both UG and PG students, considering the relative strengths of the drivers and whether they vary across these two groups.

---

[1] Research on field differences was pioneered in the 1980s. Even at that time, Hearn (1985) argued that the area "has been too little explored" (Hearn, 1985: 430).

Impact of useful teaching on module satisfaction in UG and PG students

As noted, empirical research on student satisfaction has considered not only the statistical significance of explanatory variables but also their comparative magnitudes. This allows for identification of important variables and the relative importance of any given determinant (Broder & Dorfman, 1994; Hearn, 1985). Such research, for example, has found both a strong (i.e. comparatively large coefficient) and statistically robust relationship between survey questions gauging various aspects of teaching quality and overall course satisfaction (DeShields et al., 2005; Hearn, 1985; Krahn & Bowlby, 1997; Letcher & Neves, 2010; Thomas & Galambos, 2004). The early work of Hearn (1985), for example, looked at programme level determinants of satisfaction. He found especially strong effects "from indicators of teaching ability" (Hearn, 1985: 421). Krahn and Bowlby (1997) tried to further empirically isolate the impact of teaching by including additional control variables. They again found teaching quality important: "our study demonstrates much more conclusively that the experience of good teaching translates into greater satisfaction with the overall university experience" (Krahn & Bowlby, 1997: 171). Recent UK based studies (Kandiko and Mawer 2014, Neves and Hillman 2017 and Green et al. 2015) confirm this viewpoint. In their summary of the literature on course satisfaction Green et al (2015: 131) conclude "Teaching variables, particularly teaching quality and expertise, tend to show the strongest relationships with student satisfaction".

Within the literature looking specifically at student satisfaction in business schools, teaching quality also emerges as an important determinant). Bennett (2003), for example, looking at a UK business school, confirms the "critical importance of teaching quality as a determinant of student satisfaction" (Bennett, 2003: 137). Agnew et al. (2016) identify teaching quality as the main determinant of the National Student Survey ranking in the UK. Deshields et al. (2005), looking at US business school, finds faculty and classes as "key" factors" in influencing satisfaction (p. 137), as do Letcher and Neves (2010). This leads us to predict a strong influence of teaching on satisfaction at the module level for both UG and PG students.

> *H 1a: teaching and learning activities involving direct student-teacher contact time (i.e. in lectures and seminars) are <u>strong</u> drivers of module satisfaction for both UG and taught PG students.*

While direct contact teaching is likely to be an important determinant of satisfaction, are large class (i.e. lecture) or small class (seminar) environments better for fostering student satisfaction and are there differences across UG and PG students? Extant student satisfaction literature has not yet investigated these questions, to our

knowledge, although we note an increasing interest in active/collaborative learning spaces (see Clement 2019) and studies on the impact of school class sizes on learning (see Higgins et al 2011). We start from the premise that seminars, typically involving closer interaction between the teacher and student, are likely to be more impactful because of the interactions that can take place in small classes..

> *H1b: small class direct contact teaching interactions (i.e. in seminars) have a <u>larger</u> impact on student satisfaction with module quality than large class activities (i.e. in lectures) for both UG and PG students.*

Indirect influences on perceptions of module quality

Moving beyond the impact of direct contact teaching there are numerous factors that do not involve direct student-teacher interaction (i.e. indirect) that may potentially influence student satisfaction with module quality.

Impacts of 'helpful recommended readings' on student satisfaction

In UG and taught PG business school modules students are expected to undertake a considerable amount of independent self-study. This mostly involves reading of module related texts. We might expect the quality of suggested readings, therefore, to have a significant and large impact on overall module satisfaction levels. The student satisfaction literature, however, often identifies the more general category of learning resources as having only moderate impacts, at best, on satisfaction. Krahn and Bowlby (1997), for example, report "weak but significant net effects" for the impact of academic resources on satisfaction. Graduates, for example, "reporting more satisfaction with library/study space (beta = .09) and with classroom facilities (beta = .12) were more likely to express satisfaction with the overall university experience" (Krahn & Bowlby, 1997: 169). By contrast, however, the teaching related variable in their study had a much larger standardised beta coefficient of 0.37. These quantitative findings are confirmed by qualitative research using focus groups, which found little mention of library and IT learning resources as influencing perceptions of education quality (Hill et al. 2003).

Would we expect the importance of reading materials to vary by UG and PG groups? PG students tend to be older and more motivated but it is possible that they are not as familiar with the UK learning system (generally a greater proportion PGs are foreign nationals at UK universities). As a result they may rely or trust, to a greater

extent, the direct input of teaching staff. Unfortunately, there is little related literature that can shed light on the differential impacts of reading materials on student satisfaction in UG and PG groups.

Higher education necessarily requires considerable independent self-study. In a business school attracting highly qualified and well-motivated students, students are likely to be engaged with their modules. Contrary to Krahn and Bowlby (1997), who indirectly suggest a weak relationship between academic resources and satisfaction, we hypothesize a moderate impact of helpful readings on satisfaction.

> *H2a: helpful readings have a <u>moderate</u> positive impact on module satisfaction for both UG and PG students.*

Online content and presence

It would be surprising if virtual learning environments (VLEs) and online content did not have a significant impact on student perceptions of teaching quality as students rely heavily upon the internet to order their lives. VLEs can host lecture notes, audio and visual recordings of lectures, online quizzes, links to on-line materials and readings. VLEs can also host discussion forums and facilitate a wide variety of teaching and learning related activities. Used effectively, such VLEs can have significant impacts on the learning experience for both UG and PG students (Arbaugh, 2012; Meyer, 2014; Sinclaire, 2011).

> *H2b: the usefulness of the module VLE has a <u>strong</u> positive impact on UG and PG student satisfaction.*

The extent to which a module integrates well with other course modules

Whether a module integrates well within a course may potentially influence student satisfaction. Modules which have a close link to other course materials may be perceived as complementary and also may have an added bonus of being easier to pass (as supporting background knowledge from other modules can be bought to bear in assessments). Modules that do not to integrate with a programme of study that often receive lower student evaluations (Warwick and Ottewill, 2004). These modules include those that are peripheral to the students' career interests (for example, modules referring to 'study skills' in the title); modules which have a different epistemological frames of reference from other modules (for example statistics courses for social sciences students) or modules where students have little or no background knowledge or experience (for example research methods modules).

Is the influence of module integration likely to have a relatively large impact on module satisfaction? While module integration may be important, we suspect other indirect factors, such as usefulness of VLE materials to be of greater importance. We expect the impact of module integration to be similar for UG and PG students.

> *H2c: module integration with other modules has a <u>moderately</u> positive influence on perceptions of quality for both UG and PG students.*

Summative assessments

Studies on determinants of student *course* satisfaction are noticeably silent as regards the impact of assessment and feedback on satisfaction. Hearn (1985), for example, has no instrument to gauge impacts of assessment and feedback. Similarly, most later studies lack coverage of assessment and feedback (Broder & Dorfman, 1994). Krahn and Bowlby (1997) do have a questionnaire item on feedback ("instructors provided helpful feedback throughout courses"). Unfortunately, however, as they use factor analysis to create a single generalised "teaching environment" variable (composed of nine questions) the specific impact of feedback cannot be isolated in their results. Rientes, Li and Marsh (2015), in one of few useful studies in this area, (looking at *module* level satisfaction as well) found that assessment considerations were the second most important driver of overall learning satisfaction (Rienties et al., 2015: 13). Moreover, they found that if the module studied did not contribute to achievement of the "wider qualification aims", they were considerably more likely to report dissatisfaction.

> *H2d: having appropriate summative assessments is a <u>strong</u> determinant of both UG and PG module satisfaction.*

Feedback on coursework

The impact of feedback on course work or what we call here formative assessment feedback is not covered in the empirical studies of satisfaction determinants. In general, moreover, it is a topic that is considered to be under-researched in higher education studies (Weaver, 2006). Formative feedback should, in theory, strongly facilitate learning and thus be considered element of a university module. Business schools are going to ever greater lengths to provide useful formative feedback to students (for example, using more efficient mechanisms, such as via web based applications, audio feedback mechanisms and the like). We thus predict the impact of formative feedback for all business school students to be moderately positive.

H2e: formative feedback is a <u>moderate</u> determinant of both UG and PG module satisfaction.

Module difficulty as a determinant of satisfaction with quality

So called "instrumental" learning has been identified in the business school context (Ottewill and MacFarlane, 2003). Traditionally, students have been thought of as adopting either a deep or a surface approach to their learning (Marton & Saljo, 1976). Deep learning involves attempting to understand underlying concepts and ideas to find meaning. It implies greater intellectual engagement with a subject. Rather than simply learning for extrinsic reasons, to pass tests and gain qualifications. Deep learners are motivated by intrinsic reasons (i.e. interest in a subject and desire to find enlightenment via conceptual understanding) (Lucas and Myer 2005,). Instrumental learning on the other hand is similar to a strategic approach (Prosser & Trigwell, 1999) borrowing techniques from surface learning (Dyer & Hurd, 2016; Ottewill, 2003); but with arguably more emphasis placed on the desired grade and career outcome. As some students may study business school related subjects for extrinsic reasons, for example, to improve salary prospects, they may be more prone to adopting strategic approaches to their studies (Ottewill & McFarlane 2001). "Instrumental" learning is arguably a form of shallow or "surface" learning, utilised with the purpose of primarily passing summative assessments as efficiently as possibly (Koris, Ortenblad, & Ojala, 2016). Thus, the starting point for developing our hypotheses regarding differences in the drivers of satisfaction between business school and the wider student body, is that the former are, on the whole, more likely to be instrumental learners (Ottewill, 2003; Ottewill & MacFarlane, 2001). This, we presume, shapes their perceptions of teaching quality. We also understand that instrumental learners will often display an "antipathy towards subjects that are not self-evidently relevant or make considerable intellectual demands" (Ottewill, 2003: 189). This may shape their views when it comes to the evaluation of their taught modules.

The increased marketisation of higher education in many countries some argue, has introduced a consumer framework rationale which leads to students being identified as consumers to be satisfied (Naido and Jamieson 2005). If students see themselves as consumers, they are more likely to behave as consumers rather than as learners (Universities UK 2017). Once they seem themselves as consumers, the perception that they are getting value for money becomes an important domain in their overall construction of satisfaction (Kandiko and Mawer 2014, Neves and Hillman 2017). In this consumer paradigm, going to university becomes an investment in a

future career, rather than an opportunity to develop a deep engagement with an academic subject (Bunce at al 2016). Consumers are perhaps more prone to rate their courses based on how hard they had to work, rather than how much they learnt (Feldman 1988, Bragga et al 2014 and Stroebe 2016). We therefore expect that modules which are perceived to be harder generally receive lower overall evaluations although this may be a less noticeable phenomenon among PG students undertaking higher level degrees. We predict that they may expect to be challenged and therefore be less sensitive to the impact of difficulty.

H3a: perception of overall teaching quality is <u>negatively</u> affected by perceive module difficulty.

H3b: the negative impact of higher difficulty level is <u>greater</u> for UG students than taught PG students.

METHODOLOGY

One of the most important ways universities (in the UK and elsewhere) measure student satisfaction of teaching (and in turn inferences about teaching quality – correctly or incorrectly) is via MEQ surveys. These are given to students towards the end of a taught module. MEQ survey questions are based upon Likert scales with choices typically ranging from 'strongly disagree' (1) to 'strongly agree' (5) 'disagree' (2) or 'agree' (4) and 'neither agree or disagree' (3). Here we use the averages of all module responses from 470 UG and 93 taught PG modules of student responses to 15 questions concerning the quality of a module. We do so in the context of Business, Accounting, Finance and Economics students studying in one UK business school. We look to explore the statistical significance and magnitude of the factors that determine student satisfaction with regards to overall module quality – i.e. the final question on the MEQ (see Table 1 for MEQ the MEQ questions used in this study). We use ordinary least squares regression analysis on the pooled average module responses to student MEQs to explore what determines the overall 'satisfaction' rating of a module. The final question asked in our survey data is: "Overall, I am satisfied with the quality of this module". This is our dependent variable. The 470 UG module responses (delivered across all levels of teaching, i.e. first, second and third year) covers six years (2010-2016) totalling 21,096 responses and 93 PG Masters level modules (four years, 2012-16) with 4,429 PG responses. We pool the undergrad and PG data samples initially (as the same module may have different module convenors in different years).

Dependent variable: is the average for each module of the Likert scale responses (scaled from 1 to 5) of perceived overall satisfaction with quality, phrased as: '*Overall, I am satisfied with the quality of this module*' (i.e. question 15 in table 1).

Table 1: MEQ questions

Questions taken from the MEQ survey
1. The learning outcomes of the module were clear
2. The learning outcomes of the module were fully achieved
3. The module developed in a coherent way which increased my understanding and knowledge of the subject.
4. Readings, website, lectures etc. were well integrated
5. The virtual learning environment was helpful
6. The level of difficulty was appropriate
7. Percentage share of respondents for whom module was too easy
8. Percentage share of respondents for whom module was too difficult
9. Lectures were helpful (mean of all lecturers)
10. Seminars were helpful (mean of all tutors)
11. Recommended readings were helpful
12. Feedback on formative work was helpful
13. Summative assessment is appropriate
14. Module integrates well with other modules in the programme
15. Overall, I am satisfied with the quality of this module

Explanatory variables: All explanatory variables (with the exception of some dummy variables) are initially taken from the student MEQ survey (the first 14 questions in table 1). These cover a variety of questions on modes of learning, module assessment, feedback and the helpfulness of the lectures and seminars. Many are strongly related to overall perceived quality. Multicollinearity, however, is pervasive when using all 14 MEQ questions. This is perhaps unsurprising as many of the survey questions (and our explanatory variables) intuitively appear to be likely to be closely related to each other. For example, clear learning outcomes (question 2) are likely to be related to having a module that develops in a coherent way (question 3). The pairwise correlations point to high levels of multicollinearity among some variables (see table 2); while some econometricians argue the problem of multicollinearity is exaggerated (as it does not lead to biased coefficient estimates), for the practical purposes of understanding the relative impact of module difficulty or easiness it is important to address. This is because introducing too many collinear variables creates instability in the coefficient estimates and ascertaining a consistent picture of the magnitude as well as the significance of the coefficients in question is challenged in the presence of severe multicollinearity.

To evaluate the extent of the problem we first compute pairwise correlations and then run a fully specified model (with all 14 explanatory variables as listed in table 1) so as to compute variance inflation factors (VIF). We do so in the three different samples we use (i.e. full and restricted to >25%, >50% response rates for each module). From this we identify five variables that have consistently high variance inflation factors (VIFs) when estimated across the three different samples. These variables included: the learning outcomes of the module were clear (q1); the learning outcomes of the module were fully achieved (q2); the module developed in a coherent way which increased my understanding and knowledge of the subject (q3); readings, DUO, lectures etc. were well integrated (q4); the level of difficulty was appropriate (q.5) (see Table 1) (VIFs>10). We subsequently undertook factor analysis of these collinear variables and identified a single factor (with an Eigenvalue greater than one). However, substituting predicted values from this single factor, while reducing the number of collinear variables, still did not fully resolve the problem, as the newly created factor was itself collinear with some of the remaining explanatory variables. Alternative approaches, including creating a summed average composite index, were also explored. These approaches, however, were also unable to resolve lower multicollinearity to satisfactory levels (i.e. average VIFs<5 and stability in coefficients magnitude/signs and significance). Having explored these different remedial measures we concluded that dropping the offending variables, creating a more parsimonious model, was the most suitable option.[2] This left us with the final nine explanatory variables related to our hypotheses (see table 3), to which we tailor our original hypotheses.

The explanatory variables related to *module difficulty* are measured by the percentage share of respondents for each module who found the level of difficulty was inappropriate because it was too hard. We also capture the share of respondents who found the module too easy, so as to ascertain how easiness of module influences overall satisfaction (table 1, questions 7 and 8). As these measures use a different scale to the other explanatory variables (i.e. as they use the standard 1-5 Likert options) we standardise all variables so as to get better insights

---

[2] The first four questions are closely linked to 'helpfulness of lectures/seminars', two variables included in our final model. Incorporating these collinear variables leads to inflated standard errors making it hard to establish whether explanatory variables are statistically significant or not. The question related to 'appropriateness of difficulty level' (q5) is also captured by our measure, 'percentage share of respondents finding the module too easy or difficult'.

into the comparative magnitude of each coefficient and thus of the relative impact of each explanatory variable. A predictor is classified as having a 'strong' impact if it is significant and ranked by size in the upper third of predictors. It is classified as having a 'weak' impact if it is ranked in the lower third and moderate in the mid third of the distribution.

Year and location dummy variables (as modules are taught across more than one location) are incorporated in the final pooled model (not reported in table 3). To empirically test the significance of differences between the PG and UG groups we create composite dummy variables for each of the explanatory variables. We test for their significance using the pooled sample, following the same approach as earlier studies (Hearn, 1985).

Response biases

From a methodological point of view, as we use average responses from each module, an important question arises as to whether any selection biases exist in the average MEQ responses owing to response selection biases. Modules with lower response rates, for example, may not similarly capture the underlying student opinions in that particular module when compared to modules with higher response rates. An important methodological question thus becomes: what response rate should be used? Within the academic literature exploring MEQs and their validity, some argue that a response rate greater than 50% is generally acceptable for the purposes of accurately evaluating teaching performance. On this basis we might consider restricting our sample to all modules with response rates in excess of 50% (see Table 2). A response rate in excess of 25%, however, is often deemed satisfactory for the purposes of performance evaluation in universities. This threshold increases the number of observations considerably (for UGs, which tend to have lower response rates than PG students), though may have the disadvantage of introducing response biases (i.e. it may possibly capture only those students that have strong positive or negative feelings about a module, whereas the voice of a silent majority may be missed). [3] The results we therefore report here include that of the full sample (response rate>0), which

---

[3] The Masters average response rate was 63% (standard deviation 10) and the undergraduate response rate was lower at 41% (standard deviation 15). Increasing response rates to above 50% sees a significant fall in the number of undergraduate observations. Another approach that can be used to explore non response bias is to compare the results between high response rate modules with those of low response rates (i.e. take the top quartile and compare with the bottom) to see if there are many differences.

increases our sample size to its maximum, as well as to two further restricted samples (i.e. >25% and >50%) for the purposes of comparison.

There are, naturally, some differences in the statistical significance and magnitude of the coefficients on the explanatory variables in the different samples across the different models. On the whole, however, a number of consistent results emerge.

Likert scales and use of OLS

The question of whether the sample averages of Likert scale responses can be meaningfully employed using OLS regression analysis is contested. Ideally, of course, we would use ordered logit modelling using the 25,000 or so individual student responses. This data, however, is not publicly available. On the one hand, some argue parametric tests cannot be used on Likert scales or their averages, as the underlying responses are non-parametric, based as they are on ordinal, not interval, data (Jamieson, 2004). On the other hand, however, it has been forcefully argued that such critics misunderstand parametric testing and that OLS can be employed on Likert averages. Non-normality and skewness typical with Likert data, for example, are not an issue: parametric statistics assume normality in distribution of sample means, following the Central Limit Theorem, not the data itself. In practice, moreover, it is found Pearson correlation is 'robust with respect to skewness and non-normality' (Norman, 2010: 629). Converting ordinal data to interval data, via for example the addition of different ordinal responses (as we do) is, moreover, theoretically justifiable (Norman, 2010). Norman (2010) concludes: 'Parametric statistics can be used with Likert data, with small sample sizes, with unequal variances, and with non-normal distributions with no fear of "coming to the wrong conclusion". These findings are consistent with empirical literature dating back nearly 80 years' (Norman, 2010: 631).

While the use of OLS on Likert scales is much debated (and we are unable to summarise all of the arguments for and against here), suffice to say the approach is commonly used across a broad range of academic disciplines and there is theoretical and practical justification for it (i.e. the results are reliable). By doing so we are able to draw from a much larger student population (25,000 student responses) than any previous studies. In Business School focused studies, for example, Deshields et al. (2005) used 143 student questionnaires (years not stated, US); Letcher and Neves (2010) 352 (between 2004-2008, US undergraduates); Bennett (2003) 377 (UK

undergraduates); and Malik (2010) 240 (in Pakistan) (Malik et al., 2010). To date, therefore, in total around 1,100 student responses taken from different countries in different time periods have been used to analyse drivers of student satisfaction. By contrast, our total sample consists of around 25,000 student responses.

RESULTS AND DISCUSSION

Many of the explanatory variables are highly significant and have a strong impact on perception of overall perception of average module quality. The adjusted r-squared is high (around 0.9 in each case) and our simplified model captures much of the variation in the dependent variable (i.e. MEQ related to average of the perception of overall 'satisfaction'). Only a few of the explanatory variables register consistently as being insignificant. This includes, for example, the explanatory variable on 'recommended readings were helpful' (reading), which in most of the samples has no impact on overall satisfaction. In general, the >50% and >25% response rate samples show broadly similar results, with an exception being that the formative assignment feedback is insignificant in the >50% response rate sample.

Diagnostic and robustness tests

Our data exhibits some of the issues commonly encountered with Likert data (i.e. positive skewness). We therefore undertake a series of further additional tests. This includes, firstly, use of quantile regression analysis, suggested as one suitable approach for data with skewed distributions. Secondly, we Winsorized our data at the 5% level (to remove outliers causing skewness). All results remained basically unchanged and consistent with our original OLS estimates. Further, visual analysis of the predicted error terms (via histograms) suggests the normality assumption is met, albeit heteroscedasticity may be present. We addressed this issue by using robust standard errors as well as employing a number of other remedial approaches (i.e. logarithmic transformations), to explore the robustness of our results. Owing to the relatively large sample size we do not consider multicollinearity to be problematic to the interpretation of our results.

Omitted variables could potentially bias our estimates. The adjusted r squared in our model, however, at around 0.9, is very high: about 90% of the variance in satisfaction is explained by our explanatory variables. This is considerably higher than that found in similar previous studies, which vary between 0.4 and 0.6. While it is

possible we have omitted other important explanatory variables from our model, we think this improbable given its high overall explanatory power.

Impacts of direct contact teaching

Teaching contact hours in the form of 'helpful' lectures and seminars, involving direct student to teacher contact time, as hypothesised (h1a), is the most important determinant of student satisfaction. The combined coefficients of helpful lectures (0.46) and seminars (0.29), stood at around 0.75 (for the >50% response rate sample) for UGs. This is considerably larger than the coefficients found on the other indirect determinants of student satisfaction.

The impact of lectures appears to be greater than that of seminars. This is paradoxical, as one might expect smaller seminar groups to be more beneficial than larger lecture theatres, where the potential for interaction with the lecturer is limited. However, it should be kept in mind that seminar contact time (in the university under study) is typically less than one half that of lecture time. Taking into account the resources devoted to both lectures and seminars, therefore, it might be argued their impacts are roughly equal. This does accord with Neves and Hillman (2016) who suggest class size is not a major determinant of satisfaction (and supported by our sample, in which class size when included in the regression model is found to be insignificant). Incorporation of class size as an additional explanatory variable in our model also shows that class size is not an important determinant of overall satisfaction. This finding is consistent with Neves and Hillman's recent studies of UK student satisfaction (2016 and 2017).

Contrary to hypothesis 1b, we do find some evidence for differences in the influence of seminars on satisfaction for UG and PG students. For PG students the impact of seminars is found to be weaker. Caution, however, is required in interpreting this result. This is because PG students may typically receive fewer seminars per module than UG students. Thus, we are not comparing like with like. From this it may be incorrect to finally conclude a significant difference exists.

Indirect determinants

Helpful readings (H2a.)

It is striking that helpful readings generally have far less of an impact on overall module satisfaction than predicted (H2a predicted a moderately strong relationship). 'Useful readings' is significant *only* in the UG and combined full sample (at the 5% level) but ranked in the lower third (i.e. comparatively speaking is a 'weak' predictor of satisfaction). Surprisingly, this variable is insignificant in the PG group across *all* module response rates (see table 3). There is, moreover, a statistically significant difference in the impact of readings between PG and UG groups. This is found, however, only in the full sample (i.e. for all response rates) (table 4).

One possible explanation for our finding is that either little reading is actually undertaken or alternatively, it could be that students simply disassociate module quality from supporting reading materials and judge quality more strongly on what is directly presented to them (i.e. in lectures/seminars by lecturers in person). Perhaps more trust and certainty is placed in materials that are delivered directly (i.e. and greater belief that these materials will be covered in summative assessments). Another possibility is that some students no longer consult reading materials as a first port of call, instead preferring alternatives, such as online audio-visual learning resources. Extensive materials, many of which report tens of thousands of individual hits, are available on internet sites (like YouTube), testifying to their great relevance to current learning behaviours and strategies. Clearly, further research is clearly required to establish why helpful readings are not a very strong or significant determinant of module satisfaction and why this manifests itself most strongly in the PG group.

Table 2: Pairwise correlations for average module responses to 15 MEQ questions.

	15.	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.
15. Overall Sat	1														
1. Lning OutClear	0.91	1													
2. Lning OutAchieved	0.94	0.94	1												
3. Mod Coherent	0.94	0.92	0.95	1											
4. Mod Integration	0.88	0.86	0.88	0.88	1										
5. VLE site helpful	0.74	0.75	0.76	0.74	0.86	1									
6. Appropriate Diff	0.84	0.81	0.83	0.82	0.8	0.73	1								
7. % Too easy	-0.01	0	0.02	-0.01	0.02	0.03	0.11	1							
8. % Too hard	-0.34	-0.25	-0.29	-0.3	-0.26	-0.18	-0.47	0	1						
9. Lectures helpful	0.86	0.79	0.82	0.83	0.77	0.66	0.71	0.02	-0.2	1					
10. Seminars helpful	0.75	0.68	0.69	0.71	0.73	0.6	0.6	0.02	-0.15	0.71	1				
11. Readings helpful	0.74	0.69	0.71	0.71	0.79	0.72	0.69	0.02	-0.21	0.67	0.6	1			
12. Feedback helpful	0.73	0.69	0.7	0.69	0.71	0.64	0.64	0.04	-0.17	0.64	0.66	0.66	1		
13. Summative Approp.	0.68	0.68	0.68	0.64	0.65	0.55	0.63	0.08	-0.21	0.59	0.55	0.6	0.62	1	
14. Mod Integrates Well	0.71	0.7	0.71	0.71	0.68	0.63	0.65	0.02	-0.19	0.62	0.53	0.63	0.57	0.57	1

Table 3: 'Overall perception of teaching quality' regressed on explanatory variables (standardised).

	UG modules				Masters modules				Masters and UG combined		
Module response rate	>0	>25%	>50%		>0	>25%	>50%		>0	>25%	>50%
Number of modules	470	385	151		94	94	85		564	479	236
5.Web	0.11***	0.1***	0.1**		0.17***	0.17***	0.2		0.12***	0.11***	0.13***
7.% Resp. Too Easy	-0.038*	0.0016	0.022		-0.012	-0.012	-0.033		-0.025	0.0052	-0.00032
8.% Resp. Too Difficult	-0.17***	-0.16***	-0.093***		-0.077***	-0.077***	-0.069***		-0.13***	-0.12***	-0.074***
9.Lectures useful	0.43***	0.47***	0.46***		0.58***	0.58***	0.57***		0.44***	0.48***	0.47***
10.Seminars useful	0.16***	0.17***	0.29***		...	...	...		0.15***	0.16***	0.25***
11.Readings helpful	0.055**	0.024	-0.01		0.036	0.036	0.018		0.052**	0.022	-0.017
12.Feedback helpful	0.11***	0.11***	0.071		-0.027	-0.027	0.0071		0.11***	0.10***	0.066*
13.Summative Assess	0.084***	0.083***	0.068*		0.14**	0.14**	0.1		0.088***	0.09***	0.065**
14.Module Integration	0.15***	0.14***	0.12***		0.12*	0.12*	0.13*		0.15***	0.14***	0.13***
Adj. r squared	0.87	0.885	0.893		0.9221	0.9221	0.9294		0.8798	0.8938	0.9115

Notes: All models pass F test at 1%. Standardised coefficients reported. 'Seminars useful' dropped for Masters sample owing to high collinearity with 'Lectures useful'. *** indicates significance at 1% level, ** at 5% and * at 10%.

Table 4: significance and direction of composite dummy variables used to test differences between UG and PG Groups

	Significance of Composite dummy measuring difference between PG and UG students in three different samples (according to response rate)		
	Module response rate		
	>0	>25%	>50%
Number of modules	470	385	151
5.Web site helpful	No	No	No
7.% Resp. Too Easy	No	No	No
8.% Resp. Too Difficult	Yes*** (+)	Yes*** (+)	Yes* (+)
9.Lectures useful	No	No	No
10.Seminars useful	Yes* (-)	Yes** (-)	Yes* (-)
11.Readings helpful	Yes* (-)	No	No
12.Feedback helpful	Yes* (-)	Yes* (-)	No
13.Summative Assessment	No	No	No
14.Module Integration	No	No	No

Notes: *** indicates significance at 1% level, ** at 5% and * at 10%. Sign in parentheses indicates direction of effect. Dummy variable equals 1 if PG module.

Helpful virtual learning environment (H2b.)

As predicted in H2b, a strong factor contributing to the overall evaluations of module quality is the quality of the content supporting the module on the VLE (with coefficients ranging from 0.1 to 0.2 depending upon sample). In the combined PG and UG sample, with response rates in excess of 50%, the coefficient is 0.13. This placed it as the joint third most important predictor of satisfaction (along with module integration). The positive impact of a useful VLE is perhaps unsurprising, as daily life is strongly shaped by internet access and for study purposes the online medium is also convenient. It is of some interest, however, that the medium for delivering content (i.e. the module VLE) should be of greater importance to student satisfaction than the content itself (i.e. helpful readings). Of course, as well as delivering materials (readings, lecture notes, possibly audio-visual materials and the like) VLEs are very useful for overall module management (i.e. module syllabus and structure, assessment deadlines, requirements, lecture and timetable changes and the like). It may be that it is these elements of the

VLE that are most useful to students. Recent studies of National Student Survey data in the UK (Bennett 2018, Sutherland et al.2018) suggest that students place a high premium on well-organised modules and programmes.

Module integration (H2c)

One of the most important and stable indirect factors predicting module quality, contrary to our initial hypotheses (which predicted a moderate impact), is the extent to which the module integrates well with other modules on the programme (question 14*Error! Reference source not found.*). Its coefficient varies from 0.2 to 0.11 across PG and UG modules (**Error! Reference source not found.**) and is consistently ranked among the top third of predictors by magnitude. Its impact can therefore be considered a strong predictor. There is no difference between PG and UG students regarding the impact on satisfaction – both find module integration very important to overall satisfaction.

Teachers, of course, often have limited control over the syllabus of the module that they are teaching or convening. An important potential factor affecting module evaluations is therefore essentially outside of their own control. When evaluating teaching staff based on MEQ feedback some attention should be given to degree of module integration. If a module has a history of poor performance with respect to this particular question, a module leader may justifiably expect to suffer from an *integration discount* (or, in other cases, perhaps benefit from a *premium*). Clearly, some attention should be given to programme design to maximise student satisfaction.[4] Why might the extent to which a module integrates well with other modules in the programme influence overall satisfaction? It is possible, of course, that complementarity helps increase overall enjoyment because of deepened understanding of the materials covered and heightened depth of study, which is enjoyable in its own right (or influences quality perceptions). It is also possible that that such integration lowers the overall perceived difficulty of that module and facilitates better summative assessment performance (i.e. makes it less difficult). Clearly, there are numerous possibilities as to how module integration may influence satisfaction.

---

[4] A further interpretation is that weaker teachers find it harder to relate content to existing academic studies/modules, focusing on their research which although interesting could be rather specialised and lacking in relevance to other materials.

Summative assessment is appropriate (H2d)

In a world of ever intensifying graduate competition students wish to be fairly evaluated. If they perceive the final summative evaluation to be unfair or high risk, they will take this into account in their evaluation of the module.. Instrumental learners, it is suggested, have an "an unhealthy preoccupation with summative assessment" (Ottewill, 2003: 189). They are often characterised as being more extrinsically driven than other learners (i.e. study to get a good degree and subsequent enhanced career prospects). They typically focus on attaining qualifications as efficiently as possible, with a preference towards clear guidance during their studies. It has been suggested they may exhibit "a high degree of dependence on tutors" (Ottewill, 2003: 189). While some more able students may wish to be challenged and to be treated as mature independent learners (Lucas and Myer 2005), this is not the case for instrumental learners. Students who are studying to achieve specific outcomes rather than learning for its own sake are more likely to use selective effort in assessments, they will try to seek favour with their academic supervisors, treat the pursuit of good marks as a challenge or a game which they play to win using very calculating, quite ruthless tactics (Entwistle and Ramsden 1983). If they then perceive that their calculations have been thwarted by what they perceive to be subjective or ambiguous marking in any particular module, they will be vocally dissatisfied.

Our results show that the impact of the summative assessment, while important, is generally ranked in the lower third as a determinant (i.e. by our classification system is a relatively weak driver). However, it is still a bit surprising that a factor that generally has little involvement in day to day delivery of a module should exert the impact that it does. This is somewhat consistent with the instrumental learning perspective of business school students, namely that they are very concerned about summative assessments being fair and appropriate.

Feedback on coursework (H2e)

Contrary to H2e, formative assessment feedback is an insignificant driver of satisfaction in PG modules and a comparatively weak one in the UG sample. In general, in the cases where found significant, the magnitude of its impact is comparatively small. This is also not an important determinant of aggregate MEQ scores in 50+ response rate in the UG sample, though does become significant (at the 5% level) as we increase the sample size to incorporate lower response rates. How can we explain the very limited impact of formative feedback on

student module satisfaction? One possibility is that its volume may be quite limited. Thus, in the general scheme of a module's delivery, it may play only a marginal role in influencing satisfaction.

Does this suggest that greater amounts of feedback should be given? Or that in general such feedback, is of limited interest or concern to students? It has been argued, however, that formative feedback is often poorly understood by students (Weaver, 2006). Another, less palatable interpretation, is that formative feedback is not taken very seriously by business school students. Instead of being concerned with learning from such feedback, business school students, considered to have comparatively strong instrumental learning orientations, care more about summative assessment marks and being graded fairly.

Impact of module difficulty (H3ab)

Perceived module difficulty had a significant negative impact on perceived module quality. For all samples (broken down by different response rates in Table 3) this was true, suggesting it is a robust determinant of module quality *dissatisfaction*. The 'Too difficult' coefficient ranged from -0.16 (in the full UG sample) to -0.09 (in the 50% response rate sample). As our explanatory variables are standardised, this means a one standard deviation increase in the module difficultness variable led to a 0.16 decrease in overall student satisfaction (other things being equal) in the full sample.[5] The coefficients on 'lectures' (i.e. response to 'lectures were helpful'), by way of contrast, which had the largest influence on satisfaction, ranged from 0.43 to 0.47. Thus, in the full UG sample module difficultness was the second most important determinant of overall satisfaction (by contrast, in the 25% response rate sample it fell to third position and in the 50% plus sample it fell further, to fifth position).

In contrast to the consistently significant and large negative impact of module difficulty on perceived quality, there was less evidence to suggest that modules perceived as being too easy experienced drops in satisfaction (using a similar measure, i.e. share of respondents reporting when a module was too easy). Only for the full-sample (i.e. based on 470 UG modules) was the 'Too Easy' variable significant and negative (at the 5% level,

---

[5] 19.3% (standard deviation 14.9) of undergraduate and 19.2% (14.1) of postgraduate student respondents reported their module to be 'too difficult'.

Table 3) and this significance quickly dropped away in samples with higher response rates. Moreover, the standardised coefficient stood at -0.03 in the full sample, suggesting its impact on reducing satisfaction was actually rather modest ('too hard' remember stood at -0.16). This indicates that while making modules too easy is unlikely to actually improve MEQ feedback (and may harm it), making them too hard will certainly reduce overall satisfaction.

With regards PG students the impact of module difficulty is significant only at the 5% level and the magnitude of the coefficient is significantly smaller (ranging from -0.13 -0.074). It is the fifth largest coefficient found in our model for the 50% response rate sample and fourth for full sample (i.e. >0 response rate, thus all modules included), suggesting other factors may be more important in shaping satisfaction of teaching at the PG level than UG level. To test this we look at the coefficient of the dummy interaction term, which is also significant. The results therefore support the difference hypothesis: namely that satisfaction in PGs are less sensitive to module difficulty.

During an era in which the costs of higher education have risen considerably for UK undergraduates (in particular), is it surprising that module difficulty has become an important driver of *dissatisfaction* for many UK based students? As noted, instrumental approaches to learning are considered relatively common within the business school context. Students may increasingly be seeing their commitment to higher education as a major lifetime investment, rather than a purely intellectual pursuit. In this regard, being excessively challenged by taught modules is not what they seem to be looking for. This, of course, raises some uncomfortable yet important questions about the potential longer term impacts of over reliance on student MEQs on higher education delivery in the UK, where market forces and government policy may be promoting ever greater dependence upon them.

CONCLUSIONS

Student satisfaction of teaching has recently been found to be among the most important determinants of programme value for money in UK universities (Neves and Hillman 2016). The UK higher education sector, recognising this, and increasingly being driven by powerful market forces, has gone to considerable efforts to improve student satisfaction with teaching. It is logical, therefore, to further analyse MEQ data so as to try and

understand exactly what drives student satisfaction and whether there are differences in the types and magnitudes of drivers between different types of students (such as UG and PG groups). By doing so, we can gain better insights into how student satisfaction can be most efficiently improved, as well as the merits and demerits of using module evaluation data to gauge teaching quality. This in turn has implications for government funders and policy-makers.

Implications for business schools managers and teaching staff

Direct contact time with teaching staff in lectures and seminars, as expected, is the most important determinant of student satisfaction at both the PG and UG level. While other indirect factors also do clearly matter, providing teaching that students perceive to be useful and relevant appears fundamental to getting good MEQ evaluations. Business schools that want to really excel in student evaluations, therefore, must think carefully about how they can fully develop the potential of their teaching staff or alternatively, prioritise the recruitment of academic staff with proven teaching track records.

As well as student experiences of face to face teaching, a number of surprising results emerge with regards non-contact drivers of satisfaction. The influence of module integration in a programme, something often beyond the control of a module convenor, for example, emerges as one of the most important determinants of satisfaction. It appears that colleagues need to spend more time working together to understand how their modules contribute to the overall programme of study. They also need to spend time creating and managing well-organised and helpful VLEs. By contrast, aspects of a module one would have thought from a pedagogical perspective to have been very important, including formative feedback (for PG students) and useful readings (both PG and UG), appear to have relatively small impacts on satisfaction.

The link between class size and satisfaction is not straightforward. Our evidence suggested that seminars, despite their smaller class size, had rather similar impacts to lectures. Further, when we included class size as an additional explanatory variable, it was not found to be significant. While there are estimation challenges associated with this finding (i.e. particularly endogeneity concerns), the fact no clear and negative relationship emerges suggests causal relationships between class size and satisfaction may be more nuanced than sometimes portrayed (i.e. typically larger classes are considered intrinsically bad for student satisfaction).

Implications for government funders and exercises like the UK's Teaching Excellence Framework

The fact that difficulty level is a driver of student satisfaction/dissatisfaction should perhaps not be so surprising. Students, particularly those with more instrumental approaches to learning (i.e. commonly found in business schools we would contend), consider the significant sums they invest in a higher education degree as key lifetime investments. Good grades (i.e. at least a 2.1) are increasingly seen as a prerequisite for success in the graduate job market. Under these conditions it is perhaps surprising that perceived difficulty level does not have a stronger negative impact on overall satisfaction. Our results, after all, still show that basic teaching quality has by far the largest impact on satisfaction (i.e. helpful lectures/seminars) and that other related factors, such as VLE site suitability, module integration, essay feedback and the like are equally, if not more important. Moreover, there is some evidence that making modules 'too easy' also leads to a reduction in module satisfaction.

What are the implications of all this? These results suggest that teaching in a manner that is perceived to be helpful to students is imperative, that modules should link together coherently and be supported by helpful online resources. At the same time there may be some pressure on module convenors (and business school managers) to reduce their difficulty levels in the pursuit of higher evaluations. Any changes introduced to reduce difficulty are likely to have most impact at undergraduate level. Finally, from a Teaching Excellence Framework perspective it might also be argued that our findings on the impact of module difficulty on satisfaction call into question the wisdom of using overall evaluations as a benchmark of quality, as doing so may create incentives to make modules less challenging (i.e.it may encourage further 'dumbing down').

Limitations and further research

Interpreting our results on the impact of difficulty level can only be undertaken with care. There are methodological difficulties and obstacles involved in both working with the data and its interpretation. Reported level of difficulty cannot be fully disassociated from teaching ability (an assumption made here). It is possible that teachers that do not communicate their main points well, for example, may be more likely to be perceived as teaching 'difficult' materials (when they may not really be) – and vice versa. So the 'difficulty' of a module, in practice, may not be easily separated from the capabilities (or lack thereof) in the teacher. In this instance, what

increased difficulty actually captures is lack of teaching helpfulness, i.e. the teachers capabilities. Future research would do well to try and further probe the impacts of difficulty on student satisfaction.

There are a range of additional explanatory variables whose impact it would be interesting to consider in future. How, for example, does gender influence satisfaction? The time of day or the room in which the lecture/seminar is being given or even the weather outside on the day affect evaluations? Studies could also be broadened to look at the impact of policies introduced to improve teaching evaluations. For example, does accreditation of staff to the HEA, or their participation in teacher training programmes, have any general impact on module evaluations? There are clearly many more interesting questions for future research to address within this field. The impact of MEQs on academic practices within higher education are becoming, we would contend, ever more pronounced. So it is vital we understand their impacts. We hope this paper may spur further such analysis of the determinants of student satisfaction as well as the seen and unforeseen consequences of using MEQs to evaluate staff performance.

References

Agnew S, Cameron-Agnew T, Liu A, Walker S (2016) What business school characteristics are correlated with more favourable National Student Survey (NSS) Rankings? *International Journal of Management Education*, 14, 219-227. DOI 10.1016/j.jime.2016.05.001

Arbaugh J (2012) Online and Blended Business Education for the 21st Century: Current Research and Future Directions. *Academy of Management Learning & Education,* 748–751.

Bennett R (2003) Determinants of Undergraduate Student Drop Out Rates in a University Business Studies Department. *Journal of Further and Higher Education* 27(2): 123–141.

Bragga M, Paccagnella M and Pellizzari M (2014) Evaluating students' evaluations of professors. *Economics of Education Review*, 41, 71-88

Broder J and Dorfman J (1994) Determinants of teaching quality: What's important to students? *Research in Higher Education* 35(2): 235–249.

Bunce L, Baird A and Jones S (2016) The student-as-consumer approach in higher education and its effects on academic performance. *Studies in Higher Education.* Epub ahead of print 14 Jan 2016. DOI 10.1080/0307509.2015.1127908

Burgess A, Senior C and Moores E. (2018) A 10-year case study on the changing determinants of university student satisfaction in the UK. PLOS one 13 (2) DOI 10.1371/journal.pone.0192976

Clement N (2019) The use and impact of the collaborative lecture theatres: digging-up the foundations of the lecture theatre, University of Leeds. Available on line at: http://teachingexcellence.leeds.ac.uk/wp-content/uploads/2019/01/Martin-Clement-Final-Report-1.pdf, (accessed 2nd May 2019)

DeShields OW, Kara A and Kaynak E (2005) Determinants of business student satisfaction and retention in higher education: applying Herzberg's two-factor theory. *International Journal of Educational Management* 19(2): 128–139.

Dodeen H (2016) Student evaluations of instructors in higher education: A structural equation modeling. *Research in Higher Education Journal* 31: 1–15.

Douglas J, Douglas A, McClelland R and Davies J (2014) Understanding student satisfaction and dissatisfaction: an interpretive study in the UK higher education context. *Studies in Higher Education* 40(2): 329–349.

Dyer S and Hurd F (2016) "What's going on?" Developing reflexivity in the management classroom: from surface to deep learning and everything between. *Academy of Management Learning and Education,* 15:2, 287-303.

Entwistle N and Ramsden P (1983) *Understanding Student Learning*, Abingdon: Routledge

Feldman K (1988) Effective college teaching from the students' and faculty's view: matched or mismatched priorities? *Research in Higher Education* 28 (4): 292-329

Filak V and Sheldon K (2003) Student Psychological Need Satisfaction and College Teacher-Course Evaluations. *Educational Psychology* 23(3): 235–247.

Green H, Hood M and Neumann D (2015) Predictors of student satisfaction with university psychology courses: A review. *Psychology Learning and Teaching* 14(2): 131–146.

Hearn JC (1985) Determinants of College-Students Overall Evaluations of Their Academic Programs. *Research*

in Higher Education 23(4): 413–437.

Higgins S, Kokotsaki D and Coe R (2011) Toolkit of strategies to improve learning, Sutton Trust. Available on line at: https://www.suttontrust.com/about-us/education-endowment-foundation/teaching-learning-toolkit/ (downloaded 2nd May 2019)

Hill Y, Lomas L and MacGregor J (2003) Students' perceptions of quality in higher education. *Quality Assurance in Educatio*n 11(1): 15–20.

Kandiko C and Mawer M (2014) Student expectations and perceptions of higher education, a study of UK higher education. Kings College Londo/QAA. Available online at:

https://www.kcl.ac.uk/study/learningteaching/kli/People/Research/DL/QAAReport.pdf (accessed 29th June 2018).

Koris R, Ortenblad A and Ojala T (2016) From maintaining the status quo to promoting free thinking and inquiry: Business students' perspective on the purpose of business school teaching, Management Learning. Epub ahead of print 22 Sept. 2016. DOI: 10.1177/1350507616668480

Krahn H and Bowlby JW (1997) Good teaching and satisfied university graduates. *Canadian Journal of Higher Education* 27(2/3): 157–179.

Letcher D and Neves J (2010) Determinants of undergraduate business student satisfaction. *Research in Higher Education Journal* 6(March): 1–26.

Lucas U and Myer J (2005) Towards a mapping of the student world: the identification of variation in students' conceptions of, and motivations to learn, introductory accounting. *The British Accounting Review* 37: 177-204.

Malik M, Danish R and Usman A (2010) The impact of service quality on students' satisfaction in higher education Institutes of Punjab. *Journal of Management Research* 2(2): 1–11.

Marton F and Saljo R (1976) On qualitative differences in learning - 1: outcome and process, *British Journal of Educational Psychology* 46: 4-11.

Meyer KA (2014) Student Engagement in Online Learning: What Works and Why. *ASHE Higher Education Report* 40(6): 1–114.

Nadiri H, Kandampully J and Hussain K (2009) Students' perceptions of service quality in higher education. *Total Quality Management & Business Excellence* 20(5): 523–535.

Naidoo R and Jamieson I (2005) Empowering participants or corroding learning? Towards a research agenda on the impact of student consumerism in higher education. *Journal of Education Policy* 20(3):267-281.

Neves J and Hillman N (2016) The 2016 Student Academic Experience Survey, York, Higher Education Academy. Available onine at:

https://www.heacademy.ac.uk/system/files/student_academic_experience_survey_2016_hea-hepi_final_version_07_june_16_ws.pdf (accessed 23rd January 2017).

Neves J and Hillman N (2017) The 2017 Student Academic Experience Survey, York, Higher Education Academy. Available online at: http://www.hepi.ac.uk/wp-content/uploads/2017/06/2017-Student-Academic-Experience-Survey-Final-Report.pdf (accessed 29th June 2018).

Neumann Y and Neumann L (1981) Determinants of Students Satisfaction with Course Work - an International Comparison between 2 Universities. *Research in Higher Education* 14(4): 321–333.

Ottewill R & McFarlane B (2001) Effective Learning and Teaching in Business and Management, London: Kogan Page.

Ottewill R (2003) What's wrong with instrumental learning? The case of business and management, *Education and Training* 45(4): 189-196.

Popp N, Weight EA, Dwyer B, Morse A and Baker A (2015)Assessing Student Satisfaction Within Sport Management Master's Degree Programs. *Sport Management Education Journal* 9(1): 25–38.

Prosser M and Trigwell K (1999) *Understanding Learning and Teaching*, Buckingham: Open University Press.

Rienties B, Li N and Marsh V (2015) Modelling and Managing Student Satisfaction: Use of Student Feedback to Enhance Learning Experience Subscriber. Gloucester: Quality Assurance Agency**.** Avaiable at: http://oro.open.ac.uk/46057/7/Modelling-and-Managing-Student-Satisfaction-15.pdf (accessed 23 January 2017).

Serenko A (2011) Student satisfaction with Canadian music programmes: the application of the American Customer Satisfaction Model in higher education. *Assessment & Evaluation in Higher Education* 36(3): 281–299.

Shurden M, Santandreu J and Shurden S (2016). An application of partial least squares path analysis to student satisfaction, *Academy of Education Leadership Journal* 20(2): 51–62.

Sinclaire JK (2011) Student satisfaction with online learning: Lessons from organizational behavior. *Research in Higher Education* 11: 1–20.

Stroebe W (2016) Why good teaching evaluations may reward bad teaching, *Perspectives on Psychological Science* 11(6): 800–816

Sutherland D, Warwick, P., Anderson, J. & Learmonth, M. (2018). How do Quality of Teaching, Assessment and Feedback Drive Undergraduate Course Satisfaction in UK Business Schools? A Comparative Analysis with Non-Business School Courses using the UK National Student Survey. *Journal of Management Education* forthcoming.

Tessema M, Ready K and Yu W (2012) Factors A ffecting College Students'S atisfaction with Major Curriculum: Evidence from Nine Years of Data. *International Journal of Humanities and Social Science* 2(2): 1–44.

Thomas E and Galambos N (2004) What Satisfies Students ? Mining Student-Opinion Data with Regression and Decision Tree. *Research in Higher Education* 45(3): 251–269.

Universities UK (2017) Education, consumer rights, and maintaining trust, what students want from their university. Available online at https://www.universitiesuk.ac.uk/policy-and-analysis/reports/Documents/2017/education-consumer-rights-maintaining-trust-web.pdf (accessed 29th June 2018).

Ward J and Shortt H (2102) Evaluation in management education: A visual approach to drawing out emotion in student learning, *Management Learnin*g 44(5): 435-452

Warwick P and Ottewill R (2004) How can problem subjects be made less of a problem? *Teaching in Higher Education* 9(3): 337-347.

Weaver M (2006) Do students value feedback? Student perceptions of tutors' written responses, *Assessment & Evaluation in Higher Educatio*n 31(3): 379-394.