**Postgraduate student experiences: a multilevel approach of PTES data**

## Introduction

There has been an increasing interest in student satisfaction and engagement recently, both in the UK and internationally. In part this is undoubtedly linked to the growing prevalence of tuition funded by students themselves, and its rising cost in HE systems such as the UK and the US. This has increasingly positioned students as consumers whose satisfaction is an important systemic and organisational outcome in itself (Gruber et al, 2010). In addition student satisfaction has been seen as a correlate of or even as a precursor to learning (Carini et al, 2006). This repositioning of students as consumers has been mirrored in HE accountability regimes where governments increasingly aim to measure ‘customer satisfaction’ in the HE sector they directly or indirectly (through loans for student tuition, for example) fund. In the UK, this has led to the introduction of student satisfaction surveys, the most well-know of which is the National Student Survey of final year undergraduates.

In November 2015 a government consultation document (known as a ‘Green Paper’) (DfE, 2015) was released by the UK government that formulated proposals around:

* The introduction of a Teaching Excellence Framework that “will deliver better value for money for students, employers and taxpayers”;
* Increasing access and success in higher education participation for those from disadvantaged and under-represented groups;
* Creating a new single gateway for entry and a common system for all providers; and
* Establishing a new Office for Students to promote the student interest and ensure value for money, and to reduce the regulatory burden on the sector

Student satisfaction plays a key role in the proposals. Point 12 in the Introduction and Executive Summary of the Green Paper describes it thus: “Course quality, teaching intensity and contact hours are all examples of information that are relevant to students. Information from the National Student Survey (NSS) (involving around 300,000 final-year undergraduates each year since 2004) and the annual, Higher Education Policy Institute surveys (undertaken with Higher Education Academy in 2015), gives some insight.” (p. 11). This growing role of student satisfaction in both accountability measures and various university league tables of course supposes that HEI’s themselves are key influencers of student satisfaction, which is not necessarily borne out by the evidence (Marsh & Cheng, 2008).

Given these issues in the changing landscape of HE it seems pertinent to look more closely at the way students perceive HE in the UK, not just in terms of the widely publicised undergraduate measures, but also in the less widely studied but no less important postgraduate sector. Postgraduate taught education has benefits to the individual in terms of income and employment, and to society in terms of the development of analytical thinking, problem solving, and specialised skills (Universities UK, 2014). The total number of postgraduate taught students in the UK numbered over 420.000, though this represents a significant decrease since the high point of PGT student recruitment in 2011 when over 480.000 PGT students studied at UK universities. This makes the study of student satisfaction among postgraduate taught students timely. To this end, this paper looks at data from the Postgraduate Taught Student Experience Survey (PTES), a national survey of PGT student satisfaction, in which 100 HEI’s and almost 70000 students participated in the most recent (2015) wave.

### Defining student satisfaction

Student satisfaction is a concept ultimately derived from business marketing, and it is often seen as poorly articulated in HE (Garcia-Aracil, 2009). A commonly used definition in the field is that of Elliott and Shin (2002, p. 198), who define student satisfaction as “the favourability of a student’s subjective evaluation of the various outcomes and experiences associated with education. Student satisfaction is being shaped continually by repeated experiences in campus life”. Most commonly in this perspective, universities are seen as service providers, with satisfaction hypothesised to be both a cognitive and affective process, and derived to a large extent but not exclusively from service quality. Service quality in HE is in turn typically derived from factors such as quality of teaching, facilities, social aspects and work experience (Clemes et al., 2008). Other factors that may influence satisfaction are price, convenience and availability, though these are often not as clearly differentiated in HE as in other service industries. Satisfaction is seen as important in service industries, and in HE in particular has been found to be related to factors such as continuing at the university, recommending the university to other (potential) students, and retention on programmes (Douglass et al, 2014; Cownie, 2014). US studies point to a significant relationship between student ratings of their teaching and their grades, though the direction of causality is brought into question by the finding that this relationship is stronger when students already know their final grade (Richardson, 2005).

Satisfaction is seen as a multifaceted construct in the business literature, and this perspective has carried over into the study of student satisfaction where student satisfaction is seen as reflecting the variety of services experienced (Gruber et al, 2010). Wiers-Jensen et al (2002), for example, see student satisfaction as encompassing several subcategories including quality of teaching, quality of supervision, content and relevance of curriculum, balance between teaching and self-tuition, quality of support facilities, and quality of physical infrastructure, while Elliott & Healy (2008) discern eleven subscales, these being academic advising effectiveness, campus climate, campus life, campus support services, concern for the individual, instructional effectiveness, recruitment and financial aid effectiveness, registration effectiveness, campus safety and security, service excellence and student centeredness. In the UK, the National Student Survey also takes this multidimensional approach and aims to measure six dimensions of student satisfaction: teaching, assessment & feedback, support, organisation, resources and personal development; though Marsh & Cheng’s (2008) analysis of the 2005 and 2006 data suggested that assessment and feedback formed separate constructs. The Postgraduate Taught Experience Survey aims to measure seven dimensions: teaching and learning, engagement, assessment and feedback, dissertation or major project, organisation and management, resources and services and skills development (Soilemetzidis et al, 2014). Teaching, student support and infrastructure are common to these different frameworks, with each divided in slightly different ways, often reflecting the specific context in which the measurements are carried out (e.g. the importance of dissertation for postgraduate taught students).

**Predictors of student satisfaction**

Whether satisfaction is seen as a valuable outcome in its own right as a correlate of learning, when it becomes both a key component of external accountability and an internal driver of university policy it becomes important to look at what factors may influence it. Key questions here are whether and what student characteristics influence student satisfaction, which can provide HEI’s and policymakers with key information on what they can and cannot affect. Educational effectiveness research in other phases of education has pointed to the importance of student intake and background in determining outcomes, though the impact of schools and colleges remains significant and important (Chapman et al, 2016), but studies on student satisfaction in Higher Education which typically have taken an approach derived from Business Studies don’t necessarily take such factors into account.

The majority of studies on student satisfaction have looked at undergraduate rather than postgraduate populations. One study that did look specifically at this population focussed on postgraduate taught students from Asia at two Australian universities. Factors such as perceived quality of education and social services were found to be related to overall satisfaction, but the sample and method did not allow meaningful measurement of differences between HEI’s and there was no control for demographic characteristics or course taken (Arambewela & Hall, 2007). A follow-up study did distinguish between students from different Asian countries, and found some significant differences, with students from India generally expressing lower levels of satisfaction than those from Indonesia, China or Thailand (Arambewela & Hall, 2009). A study specifically employing the notion of service quality mentioned above examined student satisfaction of postgraduate taught business students in the UK and the US, and found overall satisfaction to be higher in the US, and predicted primarily by overall impressions of the school and the quality of education. Home students expressed higher levels of satisfaction than overseas students (Mai, 2005).

Studies on undergraduate populations have primarily attempted to look at HE factors that can influence satisfaction. Clemes et al (2008), for example, studied the relationship between service quality and student satisfaction, finding a positive correlation with a service quality construct made up of three dimensions: interaction quality (e.g. interaction with staff), physical environment quality, and outcome quality (e.g. career development). The analyses did not look at the relative impact of HEI attended, course or demographic characteristics. In general, when looking at different aspects of service quality, teaching quality and extent of student centredness are typically found to be key determinants of overall service quality and in turn satisfaction, and this more so than elements such as social climate, physical infrastructure and the quality of services from administrative staff, though the latter are also significant. Responsiveness, level of contact with and access to academic staff have also been found to be significant in a number of studies (Wiers-Jensen et al, 2002; Douglas et al, 2007; Elliott & Healey, 2008; DeShields et al, 2005; Ali et al, 2014; Douglass et al, 2014). These findings appear to be replicated across contexts, with studies in the UK, Europe, North America and Asia showing similar results.

Of course, while most satisfaction surveys have focussed on students still at university, it may also be helpful to gain an account once they have left and can make a retrospective judgement. One pan-European study of students four years after graduating showed overall high levels of satisfaction, though levels of satisfaction were somewhat lower with regards to opportunities to participate in research, supply of teaching materials, provision of work experience opportunities, testing/grading and opportunities to interact with staff outside of the classroom. Interestingly, differences between countries (11 participated) were small, but there were significant relationships with demographic characteristics, with older and male respondents expressing greater levels of satisfaction, as did those studying mathematics-related subjects and those stating they were applying their learning in their current work environment (Garcia-Aracil, 2009). As well as demographic characteristics, a few studies have looked at other individual student characteristics, DeWitz & Walsh (2002) and Arambela & Hall (2013) finding a significant relationship with self-efficacy, and the latter reporting a negative relationship between personal hedonism and satisfaction among Asian postgraduate students, though effect sizes in all cases were weak or modest.

What is notable in the papers reviewed is that hardly any use analytical methods that account for the stratified nature of their samples, with students nested in HE institutions, which makes accurate estimation of the actual institutional effects on satisfaction and engagement problematic. Typically, structural equation models are used, which in most cases do not control for either stratification or student demographic characteristics. One study that did use multilevel modelling to account for the structure of the sample was Marsh & Cheng’s (2008) study on undergraduate students in the UK, analysing data from the National Students Survey. Using multilevel modelling to look at two years of data, they found that university attended explained only 2.5% of the variance in overall student satisfaction, though institutional rankings were stable over the two years. Around 4.5% of variance was explained by discipline. Similar findings of limited HE effects are found in studies using the Course Experience Questionnaire in Australia, which likewise does not distinguish strongly between HEI’s or courses taken (Richardson, 2005).

**The Postgraduate Taught Experience Survey**

The Postgraduate Taught Experience Survey (PTES) is a national survey of PGT student satisfaction conducted in the UK. 100 HEI’s and almost 70000 students participated in the most recent (2015) survey. The instrument consists of 39 questions, designed to elicit views of students on their experiences regarding Teaching and Learning, Engagement, Assessment and Feedback, Dissertation or Major Project, Organisation and Management, Resources and Services, Skills Development and a range of socio-demographic variables, as well as motivations for choosing their course and views on the information provided by their HEI. The scales were developed using principal components methods, and show good internal consistency (Soilemetzidis et al, 2014). The profile of respondents is broadly in line with the national PGT population, though there is a slight underrepresentation of part time students, and the surveys typically get a response rate of just under 30% (Soilemetzidis et al, 2014). The items included in the scales consist of Likert style items, going from 1 (definitely disagree) to 5 (definitely agree). The seven scales are significantly correlated, with coefficients ranging from .35 (resources and assessment) to .75 (teaching and assessment). The scale developers also posit an overall satisfaction scale. This appears justified on the basis of the correlations, but has not been explicitly tested in the analyses. For more information on the structure of the instruments we refer to the PTES reports (Soilemetzidis et al, 2014), which provide detailed information on levels of student satisfaction in the different domains. For this study, we will use the data from the 2014 and 2015 PTES editions to answer three key research questions that relate to the overarching question of determinants of student satisfaction, and in particular the respective importance of individual and institutional factors:

1. Is postgraduate student satisfaction a hierarchical multidimensional construct?
2. What proportion of the variance in satisfaction is explained by student and institutional characteristics, thereby determining whether the PTES surveys can reliably distinguish between institutions, and between disciplines?
3. What is the relationship between key student and institutional variables and satisfaction of PGT students?

## Methodology

To answer the research questions described previously two approaches were adopted. To answer research question 1, a Confirmatory Factor Analysis was conducted to test the fit of a factor structure. To answer research questions 2 and 3 we conducted multilevel analyses. The following paragraphs will give a justification of these choices.

*Description of the dataset*

For these analyses we utilised the 2014 and 2015 PTES surveys. 61683 responses were recorded for PTES 2014, and 69187 responses were recorded for PTES 2015. In both cases they came from 100 HEI’s. Missing value analyses were carried out for both years. In both years missing value percentages were below 5% for all variables on all seven scales except for dissertation/major project, were they were between 30% and 45% due to respondents not all having reached the dissertation phase of their studies at the time of survey completion, and for a few variables in the resources scale where they reached 8.2% in 2014. In order to check whether imputation would be an option, Little’s MCAR test was conducted for each scale in each year, but was found to be highly significant, and therefore no imputation was employed. Missingness was not found to be systematically related to demographic student characteristics or institution.

The dependent variables for the analysis of PTES were the 6 scales for Teaching and Learning, Engagement, Assessment and Feedback, Dissertation or Major Project, Organisation and Management, Resources and Services and Skills Development, and the overall satisfaction scale which is the sum of the individual scales. The full scales and items can be found in Soilemetzidis et al (2014).

The independent variables can be divided into respondent characteristics and institutional characteristics. Respondent characteristics taken from the PTES survey were:

* Age, an ordinal variable in which ages are grouped, with 18-23 as the reference category;
* Gender, a categorical variable with Male as the reference category;
* Disability, a categorical variable with disabled (yes) as the reference category;
* Country of origin, grouped, a categorical variable with UK as the reference category;
* Ethnic background, grouped, a categorical variable with non-BME as the reference category;
* Student status (full or part time), a categorical variable with full time as the reference category.

Institutional characteristics in the PTES dataset were:

* Type of university, a categorical variable with 1994 group as the reference category
* Country, a categorical variable with Scotland as the reference category. In 2015 London was added as a separate category to this variable.

These independent variables were selected as they were the main potential predictors available in the PTES datasets.

### *Confirmatory Factor Analyses*

In order to test the validity of the proposed scales, Confirmatory Factor Analyses (CFA) were undertaken. CFA is a method used to test the fit of a hypothesised factor structure to a data matrix, and is therefore a suitable method in cases such as these where variables are hypothesised to form specific predetermined scales. CFA is a so-called ‘latent variable’ model, where responses to individual items are seen as resulting from underlying traits (e.g. agreement with the item ‘assessment arrangements in marking have been fair’ are seen as reflecting broader satisfaction with assessment). The hypothesised scales therefore form the underlying latent traits that determine responses to the individual items. CFA allows us to test our hypothesised scale structure, by looking both at overall fit of the model to the data matrix, and the relationship of the latent traits to the individual variables.

Overall model fit is tested using a range of indices which have been developed in the context of structural equation modelling (of which CFA is a part) over time. The most venerable is the Chi Square test, which tests whether or not there is a significant difference between covariances predicted by our model and those present in the actual data matrix. A non-significant Chi Square, signalling no significant difference between our model and the data, indicates good fit. However, this test (like all significance tests) is highly sensitive to sample size, such that in large samples like this one a non-significant Chi Square is not readily obtainable. As a result alternative fit indices have been designed. In this set of analyses we have used the Goodness of Fit Index (GFI), an absolute fit index, the Comparative Fit Index (CFI), as the name implies a comparative fit index, and the Root Mean Square Error of Approximation (RMSEA), to cover a range of approaches (see Byrne, 2013 for full description of the indices). GFI and CFI are said to show good fit when over .9, RMSEA when below .8 (Byrne, 2009). As the method does not allow any missing values in the data matrix, EM imputation was used to replace missing values in these analyses (though, as mentioned above, not for the full multilevel analyses reported below).

### *Multilevel models*

Data from the PTES is hierarchical in structure. Each of the *students* surveyed studies a certain *discipline, a discipline* can be said to be nested in an *institution*. This nesting suggests that two notionally identical students based at the same institution are likely to be more similar in their responses to the survey than two notionally identical students studying at two different institutions. Of course, this is an *assumption* which needs to be tested. With multilevel modelling we can explore the unique impact of student, discipline and institutional characteristics upon students’ responses to items. The unique effect of different student, course and institutional characteristics are referred to in multilevel modelling terms as ‘fixed effects’. Multilevel modelling was used to analyse the two datasets. Multilevel modelling is an adaptation of the general linear model for hierarchical datasets, which partitions the variance in the dependent variable across the relevant levels. This solves the problem of attenuation of standard errors in standard linear regression models, which results from hierarchical samples such as this where individuals within a subject or university may be more similar to one another than they would be to the population of students as a whole (Snijders & Bosker, 2012). Multilevel modelling also allows us to more accurately model the data by taking its structure into account, and provides answers to important substantive questions such as the research questions above.

The model is calculated at two levels, with the level 1 equation being:

where

 refers to the score on the dependent variable for an individual observation at Level 1 (subscript i refers to individual case, subscript j refers to the group).

 refers to the Level 1 predictor.

 refers to the intercept of the dependent variable in group j (Level 2).

 refers to the slope for the relationship in group j (Level 2) between the Level 1 predictor and the dependent variable.

refers to the random errors of prediction for the Level 1 equation (it is also sometimes referred to as ).

And the level 2 equation is given as:

 refers to the overall intercept. This is the grand mean of the scores on the dependent variable across all the groups when all the predictors are equal to 0.

 refers to the Level 2 predictor.

 refers to the overall regression coefficient, or the slope, between the dependent variable and the Level 2 predictor.

 refers to the random error component for the deviation of the intercept of a group from the overall intercept.

And in which the dependent variables are the intercepts and the slopes for the independent variables at Level 1 in the groups of Level 2.

### *Model building*

We constructed three models which were created using the previously described multilevel approach. These are three-level models (students within disciplines within universities):

1. Model 0: the null model.
2. Model 1: a model with institutional characteristics included.
3. Model 2: a model with both student and institutional characteristics included.

The full model tables report estimates with standard errors and model fit. Significant values at the p<.05 level are indicated in italics. The models are conducted with several dependent variables, being the scales that could be calculated from the data. The models were run in three phases. In phase 1 we ran an empty model (Model 0), with only a constant as predictor as a baseline model which will provide us with a partition of the variance in the outcomes to be explained at each of the three levels. In the second phase, the organisational characteristics were added to the model to test their relationship to outcomes (Model 1), while in the third phase student characteristics will were added to the models (Model 2). Data were analysed using the MLWin software programme. IGLS estimation[[1]](#footnote-1) was used, all models converged and no non-admissible parameters were generated. As the independent variables were mainly categorical and ordinal, no centering was used. In conducting the analyses we tried to adhere to the recommendations Dedrick et al. (2009) formulated with regard to multilevel analyses:

1. We try to provide a clear description of the process used to arrive at the model(s) presented. We describe how the predictors were selected (although this is partly constrained by the available variables in the HEA datasets), and how many models were examined.
2. Throughout the analyses we did not use centering, as we contend that the interpretation of regression coefficients and variance estimates will be more meaningful without.
3. We have assumed normally distributed dependent variables. Because of the nature of the data no outliers were assumed.
4. There was missing data. The section on Missing Value Analysis (MVA) explains how we dealt with it. In sum, we proceeded with list-wise deletion for model building but then checked the final model with EM-imputed data.
5. Analyses were conducted with MLwin version 2.34. We used IGLS estimation and obtained convergence on the results.
6. The results of the models, including parameter estimates and model fit, are fully reported in the result tables in the full text and in appendix A.
7. In addition, standard errors for the parameters of interest are provided, consistent with the general reporting guidelines provided by the APA Task Force on Statistical Inference (Wilkinson & Task Force on Statistical Inference, 1999). We are aware of the limitations of statistical significance tests but are confident these elements provide insight in the findings.

We used three-level multilevel models, with students nested in disciplines, and disciplines in HEI’s. The discipline classification used was the JACS-1 classification, which identified 20 discipline clusters (https://www.hediip.ac.uk/wp-content/uploads/JACS\_Report\_2013-07.pdf). As the study concerns secondary data analysis, ethical Approval from the University of Southampton was sought and obtained (number 16933).

## Results

We report the descriptive statistics, the results of the confirmatory factor analyses, and the multilevel models respectively.

### *Descriptives*

### TABLE 1 ABOUT HERE

Table 1 shows that satisfaction was similar across the two years (with overall satisfaction somewhat higher in 2015). On the subscales the highest levels of satisfaction were recorded for resources and services, the lowest for organisation and management and assessment and feedback. Table 2 shows the frequencies for the independent variables over the two years.

TABLE 2 ABOUT HERE

*Confirmatory Factor Analyses*

The models were run for five out of the six scales, as the Dissertation or Major Project factor contains too many missing values to impute missing values. The first model tested followed precisely the hypothesised 6 scales as depicted in Figure 1, with each scale treated as an underlying latent variable that predicts item responses. Each item is also associated with an error term.

FIGURE 1 ABOUT HERE

The fit indices for this model in both years are depicted in Table 3.

TABLE 3 ABOUT HERE

As can be seen in Table 3, this model showed poor fit with the data. Exploration of the path coefficients, however, indicated that this was not due to poor fit of the items with their respective factor. Rather, there appeared to be strong correlations between the latent variable, leading us to hypothesise the existence of an underlying higher order latent variable we could term overall satisfaction. This was added to the model in Figure 2. In addition, error covariances were freed for those manifest variables (question items) measuring the same construct where modification indices indicated a strong relationship.

FIGURE 2 ABOUT HERE

The fit indices for this model in both years are depicted in table 4

TABLE 4 ABOUT HERE

This model achieved far better fit than the original, showing moderate fit in 2014 (reasonable GFI, poor CFI, good RMSEA) and good fit in 2015 on GFI, CFI and RMSEA. In order to avoid capitalising on chance we did not add any error covariances suggested by the modification indices.

Therefore, in the multilevel analyses we used the 7 original scales (the Dissertation/Major Project scale, though not used in these analyses, was used in the multilevel models) but in addition calculated an overall satisfaction scale from the six scales above which will also be included.

### *Multilevel models*

Overall, as demonstrated in Table 5, the first key finding is that the vast majority of the variance is explained at the pupil level, with only about 5% of variance remaining at discipline and HEI levels. In most cases discipline explains more variance than HEI. The two institutional variables, type of university and country, are not generally very significant. Individual student characteristics matter more, explaining up to 7% of total variance (see appendix A). BME students and students without disabilities are more positive, as are most of the time part-time and male students, those from the older age groups, and those from Africa and Asia.

TABLE 5 ABOUT HERE

The findings from the 2015 PTES survey confirm those from 2014, in that again (see Table 7) the vast bulk (over 90% on all scales) of variance is attributable to differences between individual students rather than their discipline or the HEI they attended. Discipline explains more variance than HEI, with the latter never explaining more than 3%.

TABLE 6 ABOUT HERE

A difference from 2014 was that the two institutional variables, type of university and country, did explain a meaningful proportion of HEI level variance. In particular, students from Small and Specialist universities and those based in London tended to express lower levels of satisfaction. At the individual student level, BME students and students without disabilities are more positive, as are most of the time part-time and male students, those from the older age groups, and those from Africa and Asia. These findings confirm those from 2014. To summarise the influence of predictors on the dependent variables in more detail, table 9 shows significant predictors and their direction. On the whole the PTES 2013 and PTES 2015 showed similar patterns. Only the ‘overall satisfaction models are presented in detail, but tables for the other models can be found in appendix A.

TABLE 7 ABOUT HERE

The overall satisfaction scale for 2014 results confirms that the bulk of variance is explained at the individual student level (almost 95%, with discipline (3.3%) explaining somewhat more variance than HEI (2%) None of the organisational variables are significant.

The student characteristics explain 6.6% of total variance. Students without a disability, BME students and part time students are more positive, while females are less positive. Students from Africa and Asia are more positive than the UK reference group, while those from Europe, the Middle East and America are less positive. The oldest student group are more positive, while those between 26 and 30 are slightly less positive than their younger peers.

TABLE 8 ABOUT HERE

Similar results were found in the 2015 sample. Of the total variance in overall satisfaction, almost 94% was attributable to the individual student level, with 4.6% attributable to the discipline level, and 1.6% to the HEI level. The institutional variables explained a third of variance at the HEI level, which however represents only 0.5% of the total variance. There were two significant predictors: students at Small and Specialist Universities and students at London universities expressed lower levels of satisfaction. The individual level variables explained 3.4% of total variance. Older, part-time, BME and non-disabled students expressed significantly greater levels of satisfaction, as did students from Africa and Asia. Females and students from other EU countries, Australasia and North America show lower levels of satisfaction compared to their respective reference groups.

TABLE 9 ABOUT HERE

## Conclusions and discussion

The purpose of this article was to perform multilevel analyses of two years of PTES data, 2014 and 2015, to obtain insight in predictors for postgraduate student satisfaction and engagement. We need to keep in mind that overall student satisfaction was high, as Table 1 indicates. We set out to answer the following questions:

1. Is postgraduate student satisfaction a hierarchical multidimensional construct?
2. What proportion of the variance in satisfaction is explained by student and institutional characteristics, thereby determining whether the PTES surveys can reliably distinguish between institutions, and between disciplines?
3. What is the relationship between key student and institutional variables and satisfaction of PGT students?

In our conclusion we add a fourth: What are the implications of these findings for sector and institutional priorities?

Firstly, in line with prior research, we find support for the multidimensional nature of the construct of student satisfaction, and for the validity of the factor structure of PTES. In addition, we find support for the notion of a hierarchical model in which the separate factors form an overarching ‘student satisfaction’ construct.

Secondly, we find that the overwhelming majority of variance for both PTES years is at the student level, with at least 90% variance explained at that level for all sub-scales and overall scores. The rest of the variance was at discipline and institutional level, with the latter being the smallest. From both sets of PTES data we can therefore conclude that it is not possible to reliably distinguish between institutions and between disciplines, as the differences between students far exceed the differences between disciplines and institutions. This result confirms the conclusions by Cheng and Marsh (2010) with NSS data, where there also was a substantial lack of agreement among students within each university in terms of satisfaction with their educational experience.

We then turned to possible predictors of this student satisfaction. Tables 7 and 8 showed the model building process for overall student satisfaction as dependent variable, and Table 9 showed the direction of all significant predictors for all the scales for both editions of the PTES survey. Full results for all subscales are presented in appendix A.

For the 2014 PTES, student characteristics explain 6.6% of total variance in overall student satisfaction. Students without a disability, BME students and part time students are more positive, while females are less positive. Students from Africa and Asia are more positive than the UK reference group on resources and services, while those from Europe, the Middle East and America are less positive. The oldest student group are more positive, while those between 26 and 30 are slightly less positive than their younger peers. For 2015, institutional variables explained a third of variance at the HEI level, which however represents only 0.5% of the total variance. There were two significant predictors: students at Small and Specialist Universities and students at London universities expressed lower levels of satisfaction. The individual level variables explained 3.4% of total variance. Older, part-time, BME and non-disabled students expressed significantly greater levels of satisfaction, as did students from Africa and Asia. Females and students from other EU countries, Australasia and North America show lower levels of satisfaction compared to their respective reference groups.

There was overlap in both years’ significant predictors. If we only look at significant predictors for both years’ overall PGT student satisfaction, we can see that positive predictors are: being a part-time student, coming from Asia or Africa, being a BME student, having no disability, and being older than 46 years old. More negative predictors are: Coming from North America or other EU states, and being female.

This study therefore has important implications for policy and practise. Firstly, as also reported by Cheng & Marsh (2010) in relation to NSS, there is little evidence that institution, or indeed discipline, has much impact on student satisfaction. Rather, differences between students account for the vast majority of such differences. However, it has to be noted that the student variables included in PTES, which relate primarily to demographic characteristics, also do not explain much of the variance in student satisfaction, which suggests that other individual characteristics may be more relevant. For example, there were no variables relating to prior attainment or SES in the datasets. Some results are noteworthy, however. Thus it appears that students from outside the UK but that are also in rich Western constructs (US, EU) show lower levels of satisfaction, while those from more different contexts in terms of culture and resources, in particular Asia and Africa, show higher levels of satisfaction, perhaps pointing to differential prior experiences of HE leading to differential expectations of PGT study. The result with regard to BME is surprising given recent studies (Boliver, 2015) that indicate that BME students might be disadvantaged. This might be explained by the fact that BME is one big ‘container’ category. Results from other HEA surveys seem to indicate that within the category meaningful distinction might be made between sub-categories of BME (Author). However, the large sample sizes and low effect sizes should lead us to exercise caution when drawing conclusions from these findings, which brings us to the limitations of the study.

As mentioned above, due to the large sample size for the datasets involved, many of the results are statistically significant even where observed differences are very small. It should be noted that the surveys in this report do not use a random sample but adopt a census approach which attempts to survey every student in the relevant population. However, this means that, like many surveys (even those which attempt a random sample) the resulting data is vulnerable to non-response bias which is not accounted for in statistical significance testing. This also is the case for any missing data. In addition, deriving continuous variables (scales) from categorical Likert scales is not without controversy given that the ‘distance’ between categories (such as ‘definitely agree’, ‘mostly agree’ and ‘neither agree nor disagree’) cannot be assumed to be the same. Finally, some categories will necessarily concern low numbers (e.g. disabled students). It is therefore advisable to remain critical about coefficients for groups with low membership.

Notwithstanding these limitations, some key implications do arise for both research and practise, however. The first is that PTES appears to be a suitably valid measure of different dimensions of student satisfaction in terms of its dimensionality and content. What is not known from these analyses is the extent to which there is any predictory relationship with student learning, a facet that extant literature on student satisfaction does not clarify further but that is clearly a key question in light of the importance with which such indicators are treated. This is by no means a given, and there are some indicators that the relationship may be highly problematic. Cho, Baek and Cho (2015), for example, find that students with good grades tend to highly rate the teaching quality of their instructors, in comparison with those who receive relatively poor grades. They interpret this finding as a ‘psychological gift’ from students to their instructors, rather than as reflecting instructor performance. Carrell & West (2010), similarly find that student evaluations reward professors who increase achievement in the contemporaneous course being taught, not those who increase deep learning. Braga et al (2014, p.85) further question the relationship, by stating that it may in some cases be negative: “These empirical findings challenge the idea that students observe the ability of the teacher in the classroom and report it to the administration when asked in the questionnaire. A more appropriate interpretation is based on the view that good teachers are those who require their students to exert effort; students dislike it, especially the least able ones, and their evaluations reflect the utility they enjoyed from the course.” Finally, Bassett, Cleveland and Acorn (2015) suggested that students’ lack of motivation and attention potentially threaten the utility of course evaluations. If all this is the case, than HEI’s and policymakers need to be clear that by foregrounding student satisfaction measures they are essentially positioning students as consumers rather than as learners, a position which at the least deserves some serious societal and political debate.

Furthermore, these results suggest that even if one accepts student satisfaction as a suitable outcome measure, it is a highly problematic accountability measure for HEI’s due to the very low proportion of variance explained at institutional level. What would appear somewhat more appropriate if using this type of data would be to discern teaching quality and student satisfaction in light of individual student characteristics. HEI’s could be asked to explain what provisions they have in place for certain groups of students. Within all of this we must keep in mind that PGT students on the whole value UK higher education positively.

Finally, what is clear from these analyses, but also from prior research, is that we are currently lacking a strong foundation regarding those factors that do predict student satisfaction. If demographic characteristics only explain a limited amount, what individual characteristics do then primarily determine outcomes? Prior research again provides only limited information, with personal characteristics such as self-efficacy also proving of limited explanatory power (DeWitz & Walsh, 2002; Arambela & Hall, 2013), and the most widely studied predictors in terms of perceived ‘service quality’ being somewhat circular in nature in light of their close similarity to the satisfaction measures used. Here clearly is a fruitful avenue for further study into student satisfaction, albeit not one that is ever likely to tell us much about either student learning or university quality.

## References

Ali, F., Zhou, Y., Hussain, K., Nair, P. K. & Ragavan, N. A. (2016) Does higher education service quality effect student satisfaction, image and loyalty?, *Quality Assurance in Education*, 24(1), 70-94.

Arambela, R. & Hall, J. (2007) A model of student satisfaction: International postgraduate students from Asia. In Borghini, S., McGrath, M. A. & Otnes, C. (Eds.). *European Advances in Consumer Research Volume 8*. Duluth, MN : Association for Consumer Research, Pp: 129-135.

Arambela, R. & Hall, J. (2009) An empirical model of international student satisfaction, *Asia*

*Pacific Journal of Marketing and Logistics*, 21(4), 555 – 569.

Arambela, R. & Hall, J. (2007) The interactional effects of the internal and external university environment, and the influence of personal values, on satisfaction among international postgraduate students, *Studies in Higher Education*, 38(7), 972-988.

Bassett, J., Cleveland, A., Acorn, D., Nix, M., & Snyder, T. (2015) Are they paying attention? Students’ lack of motivation and attention potentially threaten the utility of course evaluations, *Assessment & Evaluation in Higher Education*, *0*(0), 1–12.

Bennet, P. & Turner, G. (2014) *PRES 2013: Results from the Postgraduate Research Experience Survey* (York, Higher Education Academy).

Bjork, R. A., Dunlosky, J., & Kornell, N. (2013) Self-regulated learning: Beliefs, techniques, and illusions, *Annual Review of Psychology*, *64*, 417-444.

Boliver, V. (2015) [Exploring ethnic inequalities in admission to Russell Group universities](https://www.dur.ac.uk/sass/staff/profile/?mode=pdetail&id=9700&sid=9700&pdetail=95908), *Sociology,* 50(2), 247-266.

Braga, M., Paccagnella, M., & Pellizzari, M. (2014) Evaluating students’ evaluations of professors, *Economics of Education Review*, *41*, 71-88.

Buckley, A. (2014) *UK engagement survey: The second pilot year* (York, The Higher Education Academy).

Byrne, B. M. (2009) *Structural Equation Modeling with AMOS. Second Edition* (New York, Routledge).

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

Carrell , S.E. & West, J.E. (2010) Does professor quality matter? Evidence from random assignment of students to professors, *Journal of Political Economy*, *118*(3), 409-432.

Chapman, C., Muijs, D., Reynolds, D., Sammons, P & Teddlie, C. (2016) *Routledge International Handbook of Educational Effectiveness and Improvement Research* (London, Taylor and Francis).

Cheng, J.H.S & Marsh, H.W. (2010) National Student Survey: are differences between universities and courses reliable and meaningful? *Oxford Review of Education*, 36(6), 693-712.

Cho, D,. Baek, W., & Cho, J. (2015) Why do good performing students highly rate their instructors? Evidence from a natural experiment, *Economics of Education Review*, 49, 172-179.

Clemes, M. D., Gan, C. E. C. & Kao, T.-H. (2008) University student satisfaction: An empirical analysis, *Journal of Marketing for Higher Education*, 17(2), 292-325.

Cownie, F. (2014) *A Study of undergraduate students’ reported commitment within higher education and the impact of that commitment on intentions to emit word-of-mouth communication.* (Unpublished doctoral thesis). (Southampton, University of Southampton)*.*

Dedrick et al. (2009) Multilevel Modeling: A review of methodological issues and applications, *Review of Educational Research*, *79*(1), 69-102

Department for Education (2015) *Fulfilling our Potential: Teaching Excellence, Social Mobility and Student Choice*. Green paper. <https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/474227/BIS-15-623-fulfilling-our-potential-teaching-excellence-social-mobility-and-student-choice.pdf> Accessed 10/1/16.

DeShields, O. W., Kara, A. & 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.

DeWitz, S. J. & Walsh, W. B. (2002) Self-Efficacy and college student satisfaction, *Journal of Career Assessment,* 10(3), 315-326.

Douglas, J. A., Douglas, A., McClelland, R. J. & Davies, J. (2015) Understanding student satisfaction and dissatisfaction: an interpretive study in the UK higher education context, *Studies in Higher Education*, 40(2), 329-349.

Elliott, K. M. & Healy, M. A. (2001) Key factors influencing student satisfaction related to recruitment and retention, *Journal of Marketing for Higher Education*, 10(4), 1-11.

Elliott, K.M. & Shin D. (2002) Student Satisfaction: an alternative approach to assessing this important concept, *Journal of Higher Education Policy and Management* 24(2), 199-209.

Enders, C. K. (2006) A primer on the use of modern missing-data methods in psychosomatic medicine research, *Psychosomatic Medicine, 68*, 427–736.

Garcia-Aracil, A. (2009) European graduates’ level of satisfaction with higher education, *Higher Education,* 57(1), 1-21.

Graham, J. W. (2009) Missing data analysis: making it work in the real world, *Annual Review of Psychology, 60*, 549–576.

Graham, J. W. & Hoffer, S. M. (2000) Multiple imputation in multivariate research, in: T. D. Little, K. U. Schnable & J. Baumert (Eds.) *Modeling longitudinal and multilevel data: practical issues, applied approaches, and specific examples* (Hillsdale NJ, Erlbaum), 201–218.

Gruber, T., Fuss, S., Voss, R., Glaeser-Zikuda, M. (2010) Examining student satisfaction with higher education services, *International Journal of Public Sector Management*, 23(2), 105 – 123.

Mai, L.-W. (2005) A comparative study between UK and US: The student satisfaction in Higher Education and its influential factors, *Journal of Marketing Management*, 21(7-8), 859-878.

Marsh, H. & Cheng, J. (2008) *National Student Survey of teaching in UK universities: Dimensionality, multilevel structure, and differentiation at the level of university and discipline – preliminary results* (York, Higher Education Academy).

Richardson, J.T.E. (2005) Instruments for obtaining student feedback: a review of the literature, *Assessment & Evaluation in Higher Education*, 30(4), 387-415.

Schertzer, C. B. & Schertzer, S. M. B. (2004) Satisfaction and retention: A conceptual model, *Journal of Marketing for Higher Education* 14(1), 79-91.

Sinharay, S., Stern, H. S. & Russell, D (2001) The use of multiple imputation for the analysis of missing data, *Psychological Methods,* 6(4), 317–329.

Snijders, T. A. B. & Bosker, R. J. (2012) *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling. Second Edition* (London: Sage).

Soilemetzidis, I. et al. (2014) *The HEPI-HEA Student Academic Experience Survey 2014 . The HEPI-HEA Student Academic Experience Survey 2014* (York, Higher Education Academy).

Trowler, V & Trowler, P (2010) *Student Engagement Executive Summary* (York, Higher Education Academy).

Universities UK (2014) *Postgraduate taught education: the funding challenge.* <http://www.universitiesuk.ac.uk/highereducation/Documents/2014/PostgraduateTaughtEducationTheFundingChallenge.pdf>. Accessed 14/2/16.

Wiers-Jenssen, J., Stensakker, B. & Grogaard, J. (2002) Student Satisfaction: Towards an empirical deconstruction of the concept, *Quality in Higher Education*, (8)2, 183-195.

Wilkinson, L. & Task Force on Statistical Inference, APA Board of Scientific Affairs. (1999) Statistical methods in psychology journals: Guidelines and explanations, *American Psychologist*, *54*(8), 594604.

1. IGLS (iterative generalized least squares) is a maximum likelihood (ML) method based on estimating the random and fixed parts of the multilevel model alternately assuming the estimates for the other part are correct. This involves iterating between two GLS model fitting steps until the estimates converge to ML estimates. [↑](#footnote-ref-1)