



City Research Online

City St George's, University of London

Citation: de Than, C. (2013). Using mirror and other internal surveys in order to improve student experience. *Learning at City Journal*, 3(1), pp. 65-78.

This is the unspecified version of the paper.

This version of the publication may differ from the final published version. To cite this item please consult the publisher's version.

Permanent repository link: <https://openaccess.city.ac.uk/id/eprint/2624/>

Copyright and Reuse: Copyright and Moral Rights remain with the author(s) and/or copyright holders. Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge, unless otherwise indicated, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way. For full details of reuse please refer to [City Research Online policy](#).

Using mirror and other internal surveys in order to improve student experience
Claire de Than, Director of Student Experience, City Law School, City University
London

Introduction

This article is the first stage of a project which considers how best to use the data collected from mirror surveys and other internal student surveys to enhance the student experience, with a subsidiary aim of thereby enhancing National Student Survey (NSS) scores. The second stage, which is underway at present, combines the theoretical basis and debate explored in this article with detailed statistical analysis of internal and external survey results, to provide a greater evidential basis for decision-making and strategic planning. The research was supported as a 2011-12 Learning Development Project, at City University London, and is intended to inform educational discussion and strategy. The interim findings discussed below are readily transferable to other disciplines and other universities.

Universities have put a great deal of effort into improving student satisfaction, but not always with measurable results. Throughout the existence of the NSS, universities have experienced significant variance between student satisfaction as represented by internal measures and the levels of satisfaction reported in the NSS. This has been the case even when the internal measures take the form of mirror surveys, i.e. surveys which mirror or closely resemble the questions on the current version of the NSS. Although general morale factors and events beyond a university's control may play a strong role in the scores, they do not necessarily explain the differences, especially where the internal questions are based on those from the NSS. Both measures may be an accurate representation of student satisfaction but measuring subtly different factors, or other influences may be operating. By examining this issue, this project aims to enable better planning for the future and the development of appropriate, tailored responses to issues. The interim findings reflect examples of best practice and next steps for the strategic use of such data, including free-text comments.

The problem Part 1: usefulness of internal surveys as a method of predicting and improving student satisfaction as measured by the NSS

"I would never rate a tutor badly after being taught by them all year and building up a relationship with them since that would be a breach of trust, even if there were things they could improve" [final year student]

Student evaluation of university experience dates back to at least the 1920s, when student surveys about teaching were introduced in some universities in the United States,¹ although it was of course not until much more recently that they became common in the US and the UK.

¹ J. Kulik, 'Student Ratings: Validity, Utility, and Controversy' (2001) 109 *New Directions for Institutional Research* 9.

Launched in 2005, the NSS is the best-known external survey of university students in the UK. As Jacqueline Cheng and Herbert Marsh put it, 'The NSS is used to gather feedback from final-year university students about their satisfaction with their course'.² It is thus distinct from the national Postgraduate Research Experience Survey (PRES) and the Postgraduate Taught Experience Survey (PTES), which, as their names suggest, are focused upon postgraduate students.

Not surprisingly, the value and use of student evaluations has been the subject of extensive research and debate.³ A large proportion of this research focuses upon student evaluation of teaching as measured by internal surveys. However, there is a growing body of research about other types of student evaluation at university level. Many researchers in this area draw a distinction between reliability and validity when considering the usefulness of tests or surveys. As Michael Huemer puts it in explaining this distinction:

A test is said to be 'reliable' if it tends to give the same result when repeated; this indicates that it must be measuring *something*. A test is said to be 'valid' if it is measuring what it is intended to measure. E.g., a scale that always reads '5' whenever a *red* object is placed on it is 'reliable' but not 'valid' as a measure of weight.⁴

A test or survey that is valid should generally also be reliable, although it may be unreliable where the number of those surveyed or tested is so low that a few people changing their minds could produce a large variation. However, as Huemer's example above shows, it is possible for a test or survey to be reliable but invalid even where large numbers of people take this test or survey. In considering the validity of internal surveys, it is important to ask what exactly they are intended to measure. Student surveys may contain questions about specific issues that appear to be intended to measure some 'objective reality' such as the ability of their lecturers to explain the subject matter, but this does not mean that they are designed to do anything than ascertain student views. As Dennis Clayson and Debra Haley state, 'evaluations are generally designed to solicit students' opinions and perceptions'.⁵ This is particularly true of the NSS, which currently measures student agreement with a range of statements, such as, 'overall, I am satisfied with the quality of the course'. Even apparently 'objective' NSS questions such as 'feedback on my work has been prompt' should arguably be seen as soliciting student perception, rather than trying to measure 'reality': feedback may be provided at the first available opportunity, and in accordance with tight predetermined deadlines, but this does not mean that students will inevitably regard it as prompt.⁶ This feature of student surveys or evaluations is not necessarily a problem, as long

² J. Cheng and H. Marsh, 'National Student Survey: Are Differences between Universities and courses Reliable and Meaningful? (2010) 36 *Oxford Review of Education* 693, 694. As explained on the NSS website, one exception is that 'students repeating their penultimate year ... will be surveyed [in that year] (NOT when they eventually progress to their final year)' (http://www.thestudentsurvey.com/faqs/faqs_5.html).

³ D. Clayson and D. Haley, 'Are Students Telling Us the Truth? A Critical Look at the Student Evaluation of Teaching' (2011) 21 *Marketing Education Review* 101.

⁴ M. Huemer, 'Student Evaluations: A Critical Review', at <http://home.sprynet.com/~owl1/sef.htm>.

⁵ Clayson and Haley (n 3).

⁶ See R. Sproule, 'Student Evaluation of Teaching: A Methodological Critique of Conventional Practices' (2000) 50 *Education Policy Analysis Archives*, available at epaa.asu.edu/ojs/article/download/441/564.

as the purpose of the survey or test is clear: that is, that it is clear that the process is simply designed to measure student perception or satisfaction. Student satisfaction is after all an important matter in itself: it would be important to know what students think about their experiences even if NSS scores were not a significant factor in the reputation of universities. However, it must be remembered that it would be better to design surveys which aim to try to improve the provision for students, rather than just their perceptions of it. Hence to use only the NSS questions for internal surveys, unadapted and unembellished, may be a missed opportunity although it is of course an understandable attempt at predicting future NSS scores and taking appropriate action.

Some universities currently use ‘internal surveys [which] ... ask more detailed and wide-ranging questions [than found in the NSS]’⁷ in the belief that they will ‘give more detail and insight into concerns’.⁸ The surveys used by these universities may serve a useful function, but they would not necessarily be valid if their purpose were to predict student satisfaction as measured by NSS scores. As one university library put it:

We have compared the NSS results with those of the other institutional surveys (which are carried out for first and second year students) and our own library survey in 2007. The issues reported in the open responses are very similar, but the rating scores from NSS are consistently lower than some of our other surveys which use satisfaction scales. I think this results from the poor form of the NSS question, which invites people to think of adequacy rather than degrees of quality.⁹

As this quotation indicates, some universities prefer to use internal surveys that are not modelled on the NSS, believing that their own surveys are better-suited for their purposes. Such universities may be correct, but internal surveys that take a different form from the NSS are obviously not ideally-suited to predicting NSS results and thus to improving student satisfaction as it measured by the NSS. A problem with this is that, from an outside perspective, the NSS is seen as far more significant than internal surveys and directly and indirectly used by prospective students to help select where and what they should study: as the NSS website puts it, NSS ‘[r]esults are publicly available to prospective students and advisors to help make informed choices of where and what to study’,¹⁰ and from September 2012 will be ‘publicly available through “Key Information Sets” (KIS), which will be available on the web-sites of universities and colleges’. Unlike internal surveys, the NSS has a reputation as the ‘official’ source of information about student satisfaction: NSS results, rather than internal surveys, are factored into the well-known higher education league tables produced by *The Times* and *The Guardian*. Furthermore, it is also important to bear in mind that poorly-designed internal surveys may not pick up problem areas identified in the NSS. As one informant put it when discussing the NSS, staff at his institution ‘had been “slightly

⁷ T. Stanley, ‘The National Student Survey – Pain or Gain? (2009) 45 *SCONUL Focus* 144, 145, specifically discussing surveys about university libraries but making a point that could be applied generally.

⁸ *Ibid.*

⁹ *Ibid.*, quoting a response from an anonymous institution.

¹⁰ The National Student Survey, ‘2. How are the Survey Results Used?’, available at http://www.thestudentsurvey.com/faqs/faqs_2.html.

shocked” at their poor scoring [in the NSS] for learning resources as their own internal procedures had not picked this up as an issue’.¹¹

The problem Part 2: divergence between internal and external surveys even when mirror surveys are used

“I don’t need to fill in this form because I am happy with your teaching” [final year undergraduate student in response to internal mirror survey being distributed]

If internal surveys are not harmonised with the NSS, they will obviously be of limited utility in predicting and improving student satisfaction as measured by the latter survey. However, it is important to bear in mind that divergence between the results of internal and external surveys may occur even when universities use surveys designed to mirror external surveys such as the NSS. There may be a variety of reasons for this. First, students who complete internal surveys may doubt that they are truly anonymous, and thus be reluctant to provide negative responses because they are worried that they may be victimised by the staff or institution concerned for doing so. In the words of Thomas Whelan, a psychologist, ‘In organizational contexts, respondents may be more likely to perceive a risk associated with responding to a survey ... which when compounded by privacy concerns, might lead to unfavourable factors such as higher levels of active nonresponse or the use of response biases’.¹² In this respect, what matters is whether students have concerns about anonymity, not whether internal surveys are genuinely conducted anonymously. ‘Privacy assurances can only have the desired effect on truthfulness if respondents believe the assurances are true—if respondents have reason to believe that their privacy either is or could be compromised, they might distort their responses’.¹³ Privacy concerns are generally much less likely to be an issue with the NSS; it is well-known that the NSS is an external survey conducted by an organisation that has a legal obligation to ensure that confidentiality of personal student information is maintained, and thus respondents are generally less likely to believe that their privacy could be compromised. Thus, there may be a factor that generally mitigates in favour of higher results in internal surveys. However, it should be noted that this issue is not as straightforward as it might appear, since the environment in which a survey is conducted can distort the results. Physical environment may lead respondents to give untruthful or erratic answers. As Whelan puts it in discussing web-based surveys, ‘the relationship of a respondent’s environment to his or her perceptions of anonymity suggests that variations in respondent surroundings may possibly lead to bias in the data quality of Web-based surveys even though these variations have little to do with characteristics of the survey itself’.¹⁴ Thus, the conditions under which students complete the NSS may be significant; it may matter ‘whether or not they were alone, their perception of the possibility of others seeing their responses while taking the survey, and whether or not they were at home

¹¹ J. Machell and M. Saunders, ‘An Exploratory Evaluation of the use of the National Student Survey (NSS) Results Dissemination Website’ [January 2007] The Higher Education Academy, available at http://www.heacademy.ac.uk/assets/documents/nss/web0274_national_student_survey_full_report.pdf.

¹² T. Whelan, (2008, April). ‘Antecedents of Anonymity Perceptions in Web-based surveys’. Paper presented at the 23rd annual meeting of the Society for Industrial and Organizational Psychology, San Francisco, CA, available at http://www4.ncsu.edu/~tjwhelan/whelan_SIOP08.pdf.

¹³ Ibid.

¹⁴ T. Whelan (n 12).

when they completed the survey'.¹⁵ This is a basic concern, since some universities may reserve university computer facilities specifically for their students to complete the NSS, and thus unintentionally distort the results of that survey, especially if internal surveys are taken in different conditions. Secondly, the results of surveys may vary depending upon the mood of the students concerned, which may in turn vary throughout the course of their degree depending upon their grades and their perceptions of their likely future degree classification. There is research to suggest that 'instructors can "buy" evaluations with grades'¹⁶ in the sense that students who perceive their grades as changing often change their evaluations 'in the same direction as their perceived grade change'.¹⁷ The authors of this research conclude: 'Besides the obvious problem with validity with the grade–evaluation effect, another problem is raised by these findings. Since the evaluations change regularly over the course of the term, which one is real?'¹⁸ This research finding about the relationship between expected grades and student survey scores is of course highly controversial: there is fierce debate about whether there is a relationship between the grades students receive and the ratings students give lecturers, and, if so, the nature of such a relationship.¹⁹ However, if this conclusion about a relationship between student grades and the ratings students give lecturers is correct, it may be that there is a similar relationship between likely future degree classification, as perceived by students, and the evaluations of institutions. This relationship could lead to variances on some programmes between student satisfaction as represented by internal measures and the levels of satisfaction reported in the NSS, since the NSS is targeted at final year students whereas internal mirror surveys are usually targeted at first and second year undergraduates.

A third factor that may produce a divergence between the results of internal surveys and the NSS also relates to the relative timing of such surveys. If programmes are organised in such a manner as to produce higher levels of student satisfaction in the first and second years of three year programmes, then there is of course likely to be a divergence between student satisfaction as measured by internal surveys targeted at first and second year undergraduates and satisfaction as measured by the NSS, targeted at final year students. This may occur, for example, because of the relative perceived quality of teaching: let us suppose that a group of students encounter teaching that they regard as unsatisfactory in their third year but were satisfied with the quality of their first and second year teaching. This will be reflected in the NSS but not in internal surveys conducted in their first and second years of study. It may not even be reflected in module level surveys if students perceive the third year lecturers in question as 'good lecturers' who are not being provided with adequate guidance, training, resources or time to provide adequate support for students (there is evidence that students sometimes purposely inflate evaluations beyond what they think the lecturer deserved because they like the lecturer).²⁰

¹⁵ Ibid.

¹⁶ D. Clayson, T. Frost, and M. Sheffet, 'Grades and the Student Evaluation of Instruction: A Test of the Reciprocity Effect' (2006) 5 *Academy of Management Learning & Education* 52, 64.

¹⁷ Ibid.

¹⁸ Ibid.

¹⁹ Ibid, 53.

²⁰ See e.g. D. Clayson, 'Within-Class Variability in Student-Teacher Evaluations: Example and Problems' (2005) 3 *Decision Sciences Journal of Innovative Education* 109.

The potential impact of timing should not be ignored. It is important to remember that the NSS is targeted at students in their final term of undergraduate teaching. These students are particularly likely to be worried about impending assessments, their degree classification and their future. Students who are particularly worried are less likely to be happy, and this may be reflected in their responses to questions about their satisfaction with their programme. As psychologist Peter Frey cautions: 'it is essential that attention be accorded to the data collection process ... Collecting ratings in class during the last week of the term may tap an unrepresentative sample in an environment which is unduly influenced by an impending final exam'.²¹ While the NSS does not of course operate in precisely this manner, it arguably pays insufficient attention to the data collection process in targeting students who are particularly likely to be concerned about impending exams.

A fourth factor that may produce a divergence between the results of internal surveys and the NSS relates to the fact that university staff are forbidden from pointing out the potential long-term negative consequences for universities, programmes, and staff of poor NSS scores. In comparison, they are not necessarily prevented from doing so with internal surveys, so students may have a better appreciation of the impact of such surveys and the responsibility that is involved with completing such surveys. While students should of course respond honestly, they should do so in a responsible manner, bearing in mind that their scores and comments may have serious repercussions for those concerned.

Universities may unintentionally encourage particularly thoughtless responses to the NSS by providing particular incentives for this survey only, or extra incentives in comparison to other surveys, designed to boost response rates. As Frey says, surveys should be designed to elicit 'thoughtful and unbiased ratings'.²² It seems reasonable to conclude that students who only complete the NSS because of the hope of winning an incentive are much less likely than other students to produce thoughtful responses; after all, they were not motivated enough to complete the survey without the thought of personal reward, and their chances of personal reward are not depending upon them producing *thoughtful* responses! Thus, such incentive schemes arguably decrease the reliability and validity of the NSS, if they have an effect at all (they may not be effective in increasing response rates²³). It might be objected that the vast majority of students would take the NSS seriously, given its significance, and that students do not generally provide thoughtless responses in evaluations. However, as Clayson and Haley point out, [s]ome evidence suggests that a majority of students will evaluate anything they are asked to judge, whether or not they have any basis for the judgment'.²⁴ For instance, David Reynolds has reported that when 1,000 psychology students were asked to evaluate a speaker, 80% did so, with this speaker overall being rated as worse than six of the other lecturers, but better than three. This might sound

²¹ P. Frey, 'Validity of Student Instructional Ratings: Does Timing Matter?' (1976) 47 *The Journal of Higher Education* 327, 335.

²² *Ibid.*

²³ See P. Sturridge, 'The National Student Survey Three Years On: What Have We Learned?', available at http://www.heacademy.ac.uk/assets/documents/research/surveys/nss/NSS_three_years_on_sturridge_02.06.09.pdf, concluding: 'it is likely that the effect of incentives, per se, is small and not statistically significant; however, it is possible that some particular types of incentives may be effective in improving response rates. In order to assess this, further research that involves working more directly with students (perhaps via focus groups) would be required'.

²⁴ Clayson and Haley (n 3).

unremarkable, but 'the speaker was unique ... because he never appeared. His lecture was cancelled, but the item was retained on the evaluation form!'²⁵ Not surprisingly, this alerted Reynolds and his colleagues to the fact that there might be problems with the data produced by the survey. Sarcastically, Reynolds observes that the ability of students to evaluate lectures without attending them might have great benefits:

As students become sufficiently skilled in evaluating ... lectures without being there, and some already are, then there would be no need to wait until the end of the semester to fill out evaluations. They could be completed during the first week of class while the students are still fresh and alert. Once the evaluations are analysed, the students need only attend the most positively evaluated presentations.²⁶

Finally, a divergence between the results of internal surveys and the NSS may occur because students who complete the NSS have much less incentive to give careful, considered responses, since they are unlikely to benefit from any change in response to the NSS results. In comparison, students who complete internal surveys in their first and second years of undergraduate study have much more incentive in this respect.

²⁵ D. Reynolds "Students Who Haven't Seen a Film on Sexuality and Communication Prefer It to a Lecture on the History of Psychology They Haven't Heard: Some Implications for the University," (1977) 4 *Teaching of Psychology* 82.

²⁶ *Ibid*, 83. – presumably this would apply whether or not they actually occurred!

Possible solutions and best practice from elsewhere

Although divergence between internal and external surveys may occur even when universities use surveys designed to mirror external surveys such as the NSS, internal surveys may still provide valuable information that could improve student satisfaction as measured by scores on the NSS. Surveys based upon the NSS and targeted at first and second year students could identify sources of student dissatisfaction that would otherwise only be identified in the NSS; i.e. at a relatively early stage, where a university can address them. Furthermore, the tracking of levels of student satisfaction across year groups through the use of surveys based upon the NSS in the first and second year of a third year programme could help to reveal matters such as the one discussed above relating to the perceived quality of teaching in different years of a programme. Internal surveys could also identify areas where students are particularly satisfied and so enable a university to build upon these strengths or at least not unintentionally remove them (it is important to bear in mind that improving student satisfaction is not simply a case of identifying sources of dissatisfaction, but also involves identifying factors which make and keep students satisfied!). While internal surveys, like all surveys, have their limitations, it is important to make the best use possible of them. We can obtain some useful instruction in this respect by examining practice elsewhere and guidance relating to the use of the NSS and related data. Szerenke Kovacs, Lyndsay Grant and Fiona Hyland have produced a paper on this issue for ESCalate, a body that produced and disseminated resources relating to Higher and Further Education (this body closed in 2011). As they state, '[i]nstitutions receive NSS data in a fairly "raw" format. In order to make use of it to improve their offering, they need to interpret the data to make sense of this information'.²⁷ It is common for universities to take the NSS data seriously. However, there is wide variation as to precisely how they use this data. As Kovacs et al put it, 'NSS data appears to be widely reviewed and circulated at both departmental and institutional level. There is no single approach across institutions to interpreting NSS data, with ... [universities] respondents reporting that tutors, course leaders, teaching staff, students and student representatives may be involved'.²⁸ Kovacs et al continue:

Some institutions and departments contextualise the NSS data through comparison with previous years. One interviewee described how when they compare the results over the years they highlight them in red, amber and green; red showing a decrease compared to the previous year, green illustrating an improvement and amber showing things had remained the same. Another interviewee highlighted that in some cases they not only analyse the data over the years, but benchmark their results against comparable institutions.²⁹

According to a report to the Higher Education Funding Council by the Centre for Higher Education Studies at the Institute of Education,³⁰ using the NSS for comparisons has to be performed with care. 'The NSS ... was engineered to produce valid comparisons between courses *in the same subject area* ... It was not originally designed to be used to compare

²⁷ S. Kovacs, L. Grant and F. Hyland, 'A Study of the use of the National Student Survey to Enhance the Student Experience in Education Departments', 14, available at escalate.ac.uk/downloads/8077.pdf.

²⁸ Ibid.

²⁹ Ibid.

³⁰ P. Ramsden, D. Batchelor, A. Peacock, P. Temple, and D. Watson, *Enhancing and Developing the National Student Survey*, available at http://www.hefce.ac.uk/media/hefce/content/pubs/2010/rd1210/rd12_10a.pdf.

whole institutions, to compare subject areas (such as overall satisfaction in English and in Art & Design), or to compare scores on different aspects of the student experience (assessment compared with teaching, for example).³¹ Comparisons may commonly be made in this manner by universities and other bodies, but that does not mean that they are appropriate. 'Rank ordering of institutions or courses (such a league table) are not useful because the differences between institutions and courses are largely very small indeed'.³² Furthermore, as Palmer et al explain, there are differences between scores in different subject areas, 'but this does not mean that student experiences are worse, for example, in Art & Design than in English or Accountancy. We simply do not know the reason for the differences'.³³ Thus, it would be inappropriate for a university to use either the NSS or internal surveys to make comparisons between levels of satisfaction in programmes in different subjects. Palmer et al conclude:

An institution may appropriately, with proper caution, compare its undergraduate Engineering programme with another institution by looking at the differences between the teaching scores and the assessment scores in both places, for example. It is on less solid ground if it compares the simple scores between its own programmes of Engineering and History, or takes an average of all scales in the NSS to compare its programmes in the same subject with those in another institution.³⁴

While it would be inappropriate for universities to use surveys to compare their own programmes in different subjects, this does not mean that all comparisons are inappropriate. 'Institutions and departments may trace their own performance from year to year ... or compare results with agreed internal benchmarks'.³⁵ In doing so and in using survey scores to 'prioritise their actions in enhancing students' experience',³⁶ they should bear in mind that not all factors are equal. As Kovacs et al explain:

[S]cores on the assessment and feedback scale are consistently lower than scores for teaching and learning, but this does not necessarily mean that assessment and feedback is 'worse' than teaching and learning. In fact, the scale that most closely predicts a positive response to the question 'Overall, I am satisfied with the quality of the course' is Teaching and Learning.³⁷

If universities are most interested in improving overall satisfaction with the quality of their programmes, '[t]his may suggest that at least some of the heavy emphasis on improving assessment and feedback may be better directed towards enhancing teaching and learning'.³⁸ Thus, it might be inadvisable to increase labour-intensive forms of feedback on coursework and examination performances, even if this would improve assessment and feedback scores, if the additional labour would mean that lecturers would have less time to prepare for their teaching. Similarly, students may say that they are unhappy with the level of

³¹ Ibid, 35.

³² Ibid,

³³ Ibid, 36.

³⁴ Ibid.

³⁵ Ibid.

³⁶ Ibid.

³⁷ Kovacs et al (n 27).

³⁸ Ibid.

sporting facilities, but improving these facilities would not necessarily significantly, if at all, increase overall satisfaction with the programme since, for example, they may be utterly irrelevant to some students' perceptions of what a good university experience is about, and other students might not have ever considered using the facilities until the question was asked of them. Research indicates that, if universities want to increase the overall satisfaction scores, they should prioritise teaching and learning and other significant factors, devoting resources accordingly. However, studies using the work of Frederick Herzberg as a theoretical framework suggest that assessing the relevance of factors is not necessarily a straightforward matter, and that producing student satisfaction is not necessarily the same as, or indeed related to, reducing their dissatisfaction. The net result of intensive effort attempting to improve the score on a particular question might be no more than a '3', if that question does not measure a factor which contributes to overall student satisfaction. Herzberg, an American psychologist specialising in business management, theorised that there are certain motivating factors ('motivators') in the workplace such as responsibility and challenging work, which cause job satisfaction, but which cause very little job dissatisfaction if they are absent.³⁹ He contrasted this with a separate set of 'hygiene' factors, such as job security. He argued that hygiene factors produce dissatisfaction when they are absent, or considered unacceptable, but tend to produce workers who are 'not dissatisfied'⁴⁰ rather than 'satisfied' when they are considered good or acceptable.⁴¹ 'The important point here is that Herzberg's theory did not define satisfaction and dissatisfaction as being at opposite ends of the same continuum. The opposite of satisfaction is not dissatisfaction, but no satisfaction. The opposite of dissatisfaction is not satisfaction, but no dissatisfaction'.⁴² Adopting this framework, theorists have tried to apply its insights to students, rather than workers. For example, de Shields et al have stated:

³⁹ F. Herzberg, 'The Motivation-Hygiene Concept and Problems of Manpower' [January-February 1964] *Personnel Administration* 3.

⁴⁰ Hence the equivalent of a '3' on the NSS or mirror surveys

⁴¹ *Ibid.*

⁴² O. De Shields, A. Kara, and E. Kaynak, 'Determinant of Business Student Satisfaction and Retention in Higher Education: Applying Herzberg's Two-Factor Theory' (2005) 19 *International Journal of Educational Management* 128, 132.

In applying Herzberg's theory to this study, faculty performance and classes are directly related to the outcome from a college experience and may be considered motivators or satisfiers ... On the other hand, the performance of advising staff may be considered similar to hygiene factors or dissatisfiers that may cause dissatisfaction but not satisfaction ... While the absence of good advising staff performance may lead to dissatisfaction, its presence may not lead to satisfaction, since students may not see it as directly related to the expected outcomes from a college experience, and usually faculty also provides similar academic/career advising.⁴³

Universities and programmes within them might profitably examine whether their efforts are aimed at removing dissatisfaction, improving satisfaction, or both; it may be that in some cases efforts are being concentrated on factors which could never increase student satisfaction however much work was done.

While Herzberg's theory may provide a conceptual framework with which to analyse the significance of factors, it must be remembered that improving student satisfaction and reducing student dissatisfaction is not necessarily the same as improving student educational experience, and that universities have obligations to provide education, rather than student satisfaction. In turn, student satisfaction is an active process in which students have to be involved. Speaking as the then Chief Executive of the Higher Education Academy, Paul Ramsden pointed out in 2007: 'students do not have a "right" to be satisfied. They are themselves part of the experience ... Effort is needed to convert the opportunity into the outcome. In the last analysis, students decide their own destinies and we can only add or subtract value at the margins'.⁴⁴ There is a need to balance the demands of student satisfaction against the requirements of professional bodies and expectations of future employers, as well as the needs and resources of the institution concerned (it is important to bear in mind that many universities are research-active and thus need to devote time and resources to research and other activities). Responses to surveys which evaluate staff or institutional performance are only meaningful if we know what the students concerned value and expect, otherwise we may just be collecting evidence that spoon-feeding is valued, that compulsory core subjects are less popular than options or that tutors who expect more preparation from their students are not as liked as others. There may be variations between programmes or particular year groups about what students regard as important. Moreover, there may be variations between what is valued by students and what is important for their educational benefit. When this is the case, student expectations should be managed or challenged, rather than met. Indeed, it is arguable that one of the purposes of student surveys should be to help to identify which student expectations need challenging.

⁴³ Ibid.

⁴⁴ P. Ramsden 'Inspiring Tomorrow's Students', available at <http://www.heacademy.ac.uk/assets/documents/news/InspiringTomorrowsStudents.pdf>.

One way to determine what is valued and expected by students is to analyse the scores produced by surveys; another way is to analyse the qualitative data, or comments, provided by students in response to survey questions. The NSS allows students to comment upon any aspects of their programme that they regard as particularly negative or positive aspects of their programme. As Kovacs et al state, institutions hold 'different attitudes about the usefulness of this data and how easy it was to interpret'.⁴⁵ Some institutions take the analysis of the free-text comments seriously, believing that '[t]he feedback which students offer in this section can offer insights into the quantitative results by placing student responses within context'.⁴⁶ One such institution, Brighton University, believes that its internal process of 'categorising respondents' free-text comments and deriving frequency tables ... [is leading to] a clearer picture of students' satisfaction and dissatisfaction'.⁴⁷ Some institutions even analyse free-text responses 'more carefully than the quantitative data, seeing them as "much more illuminative and useful than the quantitative data"'.⁴⁸ In contrast, others find 'it difficult to know how much weight to give to this data as it was often contradictory and common themes were not identified'.⁴⁹ The use of text-analysis software may make it easier to identify common themes, particularly where there are many comments. Such software is already used in Australia. The Australian equivalent of the NSS is known as the Course Experience Questionnaire, or CEQ. 'Special software (CEQuery) has been developed to facilitate institutional analysis of the comments in the CEQ, which includes two similar open response items'.⁵⁰ A report for the Centre for Higher Education Studies at the Institute of Education recommends that UK institutions 'might consider using this system, or a similar one, to help them analyse the NSS open comments efficiently'.⁵¹ Of course, any such software could equally be used to help analyse free-text responses to internal surveys. However, it is vital to analyse open comments carefully. In discussing the NSS in relation to the University of Arts, London, Duna Sabri argues that '[c]ounting the qualitative data is superfluous, and of questionable value'.⁵² She continues: 'it is possible to say that there were 634 comments about assessment. This statement has very little meaning because it does not tell us anything about all those who did not use the "comments" space to mention assessment. ... The first 22 questions in the NSS give much more reliable and extensive quantitative data, which include a validated 5-item scale on assessment'.⁵³ While Sabri is correct to state that free-text comments need to be considered carefully, she goes too far in stating that counting the qualitative data is superfluous. The fact that a student identifies an aspect of a programme as particularly positive or negative does not necessarily indicate that s/he regards this aspect of the programme as particularly important: it may simply reflect the timing of the survey, with the student focusing upon recent matters

⁴⁵ Kovacs et al (n 27).

⁴⁶ University of Brighton, Academic Board, Academic Standards Committee, 'National Student Survey 2011: overview of core results and areas for priority action', available at <http://staffcentral.brighton.ac.uk/nss/2012%20files/ASC11-63%20NSS.pdf>.

⁴⁷ Ibid.

⁴⁸ Kovacs et al (n 27).

⁴⁹ Ibid.

⁵⁰ P. Ramsden et al (n 30).

⁵¹ Ibid.

⁵² D. Sabri, 'National Student Survey at the University of the Arts, London: Voice, Interpretation and Context, available at <http://www.arts.ac.uk/media/academicaffairs/asdc/asdc36/Item%2013.2%20Report%20-%20NSS%20at%20UAL%20Voice%20Interpretation%20-%20Context%20-%20Final.pdf>.

⁵³ Ibid.

that are fresh in his or her mind. Nonetheless, it does reveal that the student regarded it as important enough to be worthy of comment; that is, that the student regarded it as significant. Thus, it seems reasonable to conclude that grouping and counting comments by themes could 'help various constituencies better understand students' perspectives of their university experience'.⁵⁴ As Tony Chambers puts it, 'areas where most students provided comments of concern suggest where improvements would enable them to have a more enhanced university life experience'.⁵⁵ One of the limitations of the first 22 questions in the NSS is that they reveal what students think, but not why they think it. In so far as free-text comments provide common explanations for why students are satisfied or dissatisfied with the matters covered in the first 22 questions in the NSS and other surveys, then it may be helpful to categorise and count these explanations. It may be possible to use comments to identify matters of particular concern as well as the issues that emerge most frequently in relation to these matters. Commenting upon the responses to the United States and Canadian equivalent of the NSS at a particular institution, Chambers concludes: 'findings from this study indicate that Academic Experience presented the most significant concern for students, although the details, of these concerns may vary in each College/Faculty. Regarding students' Academic Experience comments, the issues that emerged most frequently included academic support, class size, limited course offerings, teaching assistants, teacher quality, and program expectations'.⁵⁶ The NSS concerns 22 specific questions, but it leaves a lot of ground uncovered that could be filled by similar analysis of free-text comments- most of Chambers' list is excluded from the 22 questions. The first question in the NSS is whether 'staff are good at explaining things'; this is currently the only question in the NSS about the ability of staff to provide adequate guidance about the subject matter of the programme. If an institution receives a low score in response to this question, either in the NSS or in internal survey based upon the NSS, analysing free-text comments may reveal why this is so: categorising and counting these comments may reveal that students think that certain lecturers have adequate facilities to perform adequately but simply lack the ability or dedication to do so. On the other hand, they may reveal that students think that they are not receiving adequate explanations because of large group sizes, limited contact hours, or other factors.

Conclusions

The analysis of all data gathered from surveys need to be undertaken with care, whether that data is quantitative or qualitative. It should be used in conjunction with other appropriate data, such as peer review, feedback from internal student committees and research on student experience produced by bodies such as the National Union of Students (NUS).⁵⁷ Evidence and practice elsewhere indicates that analysis of free-text comments on internal surveys could be extremely useful in predicting and preventing NSS issues at an early stage, improving student satisfaction and engagement, and in preventing attrition of student

⁵⁴ T. Chambers, 'What I Hear You Saying is ...: Analysis of Student Comments from the NSSE' (2010) 44 *College Student Journal* 3, discussing the National Survey of Student Engagement, the United States and Canadian equivalent of the NSS, but making a point that is equally applicable to the NSS.

⁵⁵ *Ibid.*

⁵⁶ *Ibid.*

⁵⁷ The NUS has surveyed 5,000 higher education students in order to produce a series of four reports about teaching and learning. The first of these reports, *Student Experience Research 2012: Part 1: Teaching and Learning*, is now available at <http://www.qaa.ac.uk/Publications/InformationAndGuidance/Documents/Student-Experience-Research-2012-Part-1.pdf>.

numbers through withdrawal and failure. The software available means that the cost and human time required to conduct such analysis may be greatly reduced. Some universities, such as the University of Hertfordshire,⁵⁸ have had surprisingly useful results from the analysis of free-text comments. Thus, such analysis appears to be worth undertaking to supplement the analysis of quantitative data, providing that, like the analysis of quantitative data, it is done carefully.

Although the NSS has a reputation as the 'official' source of information about student satisfaction, it must be borne in mind that internal surveys can be used to ask more detailed or different questions in order to give more information about matters of concern to individual institutions, such as the average number of hours each student spends studying each week throughout the academic year. It is possible for institutions to customise the NSS to their own purposes to some extent by choosing from a set of additional, optional questions and by formulating two questions which are not part of this set. However, while this is potentially useful, there is limited scope for customisation of the NSS: institutions are not allowed to adjust the phrasing of NSS questions or to include more than two questions of their own devising. As noted above, internal surveys which take a different form from the NSS are obviously not ideally-suited to predicting NSS results and thus to improving student satisfaction as it is measured by the NSS. However, the best practice may be to use internal surveys which take the NSS format as a starting point but modify it to make it better-suited to the particular purposes of the institution concerned. This might sound contentious, but it must be remembered that the universities have an obligation to provide education, which is not identical to student satisfaction as it is measured in the NSS. 'Higher education should be a transformative process that supports the development of graduates who can make a meaningful contribution to wider society, local communities and to the economy'.⁵⁹ Students who are offered a high quality education should be satisfied with it, but this cannot be guaranteed.

⁵⁸ See C. Rendell and H. Barefoot, 'Using the NSS to Inform Strategic Enhancement at University Level', available at http://www.heacademy.ac.uk/assets/documents/postgraduate/Session_9_Helen_Barefoot_and_Catherine_Rendell_final.pdf.

⁵⁹ C. Mahoney, foreword, in G. Gibbs, *Dimensions of Equality*, 2, available at http://www.heacademy.ac.uk/assets/documents/evidence_informed_practice/Dimensions_of_Quality.pdf.