

# COMPETITION IN THE TRAINING MARKET

EDITORS

TOM KARMEL, FRANCESCA BEDDIE, SUSAN DAWE

*Papers from the Competition in the Training Market Round Table  
March 2009*



Australian Government

Department of Education, Employment  
and Workplace Relations

# Contents

---

Contributors	6
Overview, <i>Tom Karmel</i>	9
Markets and central planning in meeting labour market needs: Lessons from higher education, <i>Andrew Norton</i>	18
Discussant: <i>Leslie Loble</i>	33
A market for vocational education and training in the Australian Federation, <i>Mark Burford</i>	40
Discussant: <i>Pat Forward</i>	49
Competition policy and the VET sector, <i>Richard Dennis</i>	54
Discussant: <i>Megan Kirchner</i>	62
Contestability, information asymmetry and quality signals in a competitive training market, <i>Terri Seddon</i>	65
Discussant: <i>Robyn Tudor</i>	79
Improving information flows for users of post-secondary education, <i>Nicholas Gruen</i>	82
Discussant: <i>Gerald Burke</i>	94
Possible governance structures and autonomy of TAFE institutes, <i>Lynn Meeke</i>	97
Discussant: <i>Michael Keating</i>	113
Endnotes	115
Appendix	123
VET revenue tables	123
VET funding source tables	128

# Improving information flows for users of post-secondary education

---

NICHOLAS GRUEN<sup>43</sup>  
LATERAL ECONOMICS

---

The normal market mechanism for dealing with asymmetric information is reputation. When we place a deposit with a bank, or visit a doctor, we rely on the reputation of the bank and the doctor to assure the security of our deposit and the wisdom of the advice. No regulation can ensure that banks will not go broke or that doctors will make correct diagnoses, and regulation directed to information disclosure – rules that compel banks to display their balance sheets, or require doctors to explain fully risks and prognoses – does not work well either.

(Kay 2003, p.370)

Сегодня мы работаем на репутацию. Завтра репутация будет работать на нас  
[Today we work for our reputation. Tomorrow our reputation will work for us.]

Russian saying (Picci 2007b, p.1)

---

## Introduction

In the last generation we have moved towards greater reliance on competition and private provision. Nevertheless, some naïve enthusiasms notwithstanding, a central tenet of economic orthodoxy going back to Adam Smith himself is that markets comprise an ecology of private and public goods. We ignore that proposition at our peril, as indicated by the various more and less botched liberalisations of former communist countries in the 1990s.

Thomas Friedman makes a similar point in his racy op-ed journalese:

[C]ome to Africa – it's a freshmen Republican's paradise. Yes sir, nobody in Liberia pays taxes. There's no gun control in Angola. There's no welfare as we know it in Burundi, and no big government to interfere in the market in Rwanda. But a lot of their people sure wish there were.

(Friedman 1999, p.435)

Public goods such as property rights and their enforcement are necessary even where trade occurs, as it does in local vegetable markets, in pure private goods. Much of microeconomics is devoted to analysis of circumstances in which traded private goods nevertheless take on certain qualities of public goods by virtue of various technical characteristics such as scale economies or externalities. Where these issues are too prominent to ignore, a case for public provision arises—at least in the textbook. In fact we have moved away from public provision and yet have sought to address the concerns of the textbook with hybrid public–private institutions. Thus, although we have corporatised and privatised many utilities and so moved away from public provision, where there are strong natural monopoly characteristics for nationally significant infrastructure, Australia's

competition policy regime encourages private investment, but nevertheless subjects it to various public disciplines on access to potential competitors.

The two technical characteristics of a public good are non-rivalrousness in consumption and non-excludability.<sup>44</sup> As information economists like Arrow, Akerlof and Stiglitz have pointed out, information has powerful public good characteristics. Increasingly in the age of the internet, information and media ‘content’ have these characteristics to a substantial extent. Thus the cost of an additional person downloading an ABC podcast or this year’s Budget Papers is next to nothing. At the same time, as recording companies and other publishers of content are discovering, it is increasingly difficult to allow access to one, without allowing access to others.

Further, discrete items of information *themselves* comprise both private and collective aspects. Thus, for instance, although this essay is one of many addressing the theme in this volume and in this sense is a private, potentially competitive contribution, it is written in the English language using Roman script. It would be unintelligible without these standards, which are of course pretty close to pure public goods.

If this helps us to understand why competitive markets can produce unsatisfactory results in the generation and dissemination of information, we need a dose of Hayek to remind us of the pitfalls of too crude a reliance on public provision. Hayek’s focus was not so much on the technical characteristics of discrete items of information, but on the fact that so much important information is decentralised throughout the economy.

This reasoning leads us to a presumption on which this paper is predicated. Even more than the market for those goods and services with ‘imperfections’, which require some collective involvement in their production and/or sale, the market for information requires a felicitous hybrid of private and public institutions to achieve the best possible outcomes.

## The structure of the paper

A presumption of the move to liberalise the vocational education and training sector is that consumer choice—in the first instance, student choice—is relatively well informed. For that reason, we focus on a single crucial means of informing them about the quality of the training on offer in VET institutions. This is information about evaluations which previous students of these institutions have made of the quality of their VET experience and the employment outcomes to which it has led them.

While I regard the information from student evaluations as sufficiently important for its proper design and dissemination to substantially improve VET outcomes, obviously there are many other things that students, employers, and VET administrators and teachers should be interested in, and there are many other issues that will influence the quality of VET. But for the sake of clarity and brevity, because it makes a compelling case study, because the internet has opened up fabulous new opportunities for us, and because many of the principles adopted in the case study can be generalised to other areas, this study is thus constrained.

We examine the VET Student Outcomes Survey (SOS) and the dissemination of its results.<sup>45</sup> This is compared with a very different (competitive) model of providing the public good of information about student evaluations through private sector internet review sites such as <ratemyprofessors.com>. We then examine the British universities’ National Student Survey (NSS) and the website on which its results are published <unistats.com>. This is a huge improvement on our own approach with the SOS. Still, the British approach could be substantially improved.

In elaborating such improvements I propose a new approach to generating and distributing information on student experience and employment outcomes. Fully embracing the potential of the ‘collaborative web’ or ‘Web 2.0’ and harnessing both collective and individual contributions would

improve things further again. While the example being used is from VET, the hybrid arrangements being proposed would have relevance more widely within post-secondary education, and, with some modifications, more broadly still.

## The value of student evaluation of tertiary education

A considerable research literature—most of it admittedly focused on university education rather than VET—suggests that student evaluations are extremely valuable in determining the quality of courses and teaching.<sup>46</sup> Given this, it is not surprising that, as a recent Organisation for Economic Cooperation and Development (OECD) report argues: ‘a strong quality culture may ... develop as a result of public intervention; for example, through the creation of internal quality assurance systems by TEIs [tertiary education institutions] or in response to appropriate incentives such as publishing student evaluations of their learning experience’ (Santiago et al. 2008, p.21).

Of course, like any system for evaluating the quality of educational services, student evaluations are far from the last word. Becker (2000, pp.113–15) lists six objections to using student evaluations as the sole means of evaluating university teachers. An important problem is that, as Nathan Bowling reports (2008, pp.461–2), student ratings of university professors’ teaching performance are highly contaminated by course easiness. Indeed, Bowling’s results suggested that approximately one-third of the variance in quality ratings is explained by course easiness. On the other hand, forewarned is (often) forearmed. Thus, it will often be possible to make reasonable statistical corrections for known sources of bias, such as course ease or class size<sup>47</sup> (Gillmore & Greenwald 1999), or students’ improved opinions of courses or teachers who evaluate them positively (Cruse 1987).

### The VET Student Outcomes Survey

Australia’s VET system expends considerable resources obtaining information from students on their satisfaction with their VET experience and their post-course employment status. In addition to numerous internal student surveys conducted at TAFE institutes, a national VET Student Outcomes Survey is conducted annually.<sup>48</sup> The survey is not disaggregated to the level of individual courses or teachers.<sup>49</sup> The lack of disaggregation in reporting may be for valid statistical reasons, but information disaggregated down to the individual institution, if not courses and teachers, is surely highly valuable for purposes of governance of individual institutions or to help students find courses best suited to them.

The greatest shortcoming of the survey is that the public release of details disaggregated to individual VET institutions is suppressed.<sup>50</sup>

Given the power of student evaluation to help fill out our picture of how VET courses, institutions and teachers are performing, and the low cost of doing so, the case for publishing more disaggregated results is powerful. Further, by rebuilding the system using the internet as a platform, we ought to be able to dramatically improve the current system.<sup>51</sup> And of course we could and should make it widely available to the general public, with disclosure being curtailed only for well-justified reasons, such as (legitimate) confidentiality or privacy. We now consider two other models of information provision, one profit-based, the other a modernised version of the Student Outcomes Survey, before using those examples to suggest our own improvements to Australian practice.

### Ratemyprofessors.com

Ratemyprofessors.com and similar sites offer a quite different model for generating and accessing student evaluations, funding themselves very largely from online advertisements and inviting students to rate their professors, the results of which results are then aggregated and reported on the site.<sup>52</sup> To maximise engagement and comprehensibility, the scale against which professors are measured is very simple (a rating of 1 to 5) on five different dimensions: easiness, helpfulness,

clarity, overall quality, and rater interest. A sign of the site's commercial orientation and its need to engage users is the facility given to users to click on a chilli pepper icon to score a professor as 'hot', with 'hotness' ratings also being reported, but not influencing other ratings on the quality of teaching.<sup>53</sup>

Ratemyprofessors.com provides valuable information. In one study the site's rankings provided a 0.68 correlation with student evaluations conducted by the school, with substantially higher correlations for those professors ranking highly. But with some other parameters scoring lower correlations, for instance, correlation of measures of easiness was 0.44, ratemyprofessors.com ratings should be treated with caution (Coldarchi & Kornfield 2007, p.7). It is not difficult to see why. Indeed, as in the case of Wikipedia, it's hard not to be surprised that it works at all.

There are the usual 'free riding' incentives, for example, students can access the information on others' evaluations without contributing themselves. Yet the site currently boasts 6.8 million student-generated ratings of over one million professors.<sup>54</sup> Self-selection is likely to bias both the types of people who post ratings on ratemyprofessors.com and skew their motivations. More enthusiastic students may self-select, as do some who are motivated to manipulate their professors (see Kindred & Mohammed 2005 for lurid examples). Indeed, because anonymous posting is so easy, some of the feedback will not be from students at all, but from those seeking to make mischief or even professors seeking to favourably influence their own ratings.

Ratemyprofessors.com tackles these problems as best it can.<sup>55</sup> But it is severely handicapped by its lack of power to compel, or verify the identity of students posting ratings. Some engagement from governments and/or the schools it rates would be likely to make a great deal of difference, a subject to which we turn after we examine one of the more thoroughgoing attempts to generate and disseminate student evaluations.

## Unistats.com

In January 2003 the United Kingdom Government white paper on higher education announced '[b]etter information for students including a new annual student survey and publication of summaries of external examiners' reports to help student choice drive up quality' (Department for Innovation, Universities and Skills 2003). The result was the website unistats.com, which provides comprehensive reporting of a census of final-year students—the National Student Survey (NSS).<sup>56</sup> Using a methodology not dissimilar to the Australian VET Student Outcomes Survey, the National Student Survey generates data on student satisfaction with subject areas rather than specific courses.<sup>57</sup>

Unistats allows users to obtain information at relatively fine levels of detail, for example, the user can see the breakdown of the answers from 1 (strongly disagree) to 5 (strongly agree) for each question. This will reveal differing levels of polarisation in opinion, even where averages are the same. The user can also see the scores achieved in response to specific questions. Thus, in one random search I did, the overall level of satisfaction between three pharmacy courses was similar. Yet substantial differences were evident on specific questions, like whether feedback was prompt.

<b>Disaggregating measures of satisfaction: Randomly chosen pharmacy schools</b>	<b>Agree (%)</b>	<b>Completion rate</b>
<i>Overall, I am satisfied with the quality of the course</i>		
University of Aberdeen: Medical science and pharmacy	93	42 of 54
Aberystwyth University: Agriculture and related subjects	94	52 of 74
Aston University: Pharmacology, toxicology and pharmacy	89	99 of 117
<i>Feedback on my work has been prompt</i>		
University of Aberdeen: Medical science and pharmacy	49	42 of 54
Aberystwyth University: Agriculture and related subjects	71	52 of 74
Aston University: Pharmacology, toxicology and pharmacy	60	99 of 117

Source: Results from a search on unistats.com

Receiving 4800 visits and 82 000 page impressions a day in 2007, the site is clearly proving of interest to users, many of whom are prospective students.<sup>58</sup> It is also evident that the survey means a lot to university administrators and teachers and they live in hope of improving their position on the regularly published official 'league tables' and in real fear of falling towards the bottom.<sup>59</sup> There is good anecdotal evidence that this new transparency has driven worthwhile reform, although much of the evidence appears ominously buried in managerial jargon, with liberal references to 'change management'.<sup>60</sup> It is also notable and somewhat concerning how often the response to poor results leads to initiatives to 'improve communication' with students.<sup>61</sup> Nevertheless, even though it may reflect an attempt to get students to see things in the best light for the purposes of their evaluations, this is something which most universities will seek to do, thus 'levelling the playing field' to some extent, as exam preparation does between students. And communication with students is an important part of serving their needs well. There are certainly some cases where it appears to be a sensible and important ingredient of a wider program to substantially improve services to students, with resulting outcomes. Thus the University of Manchester Dentistry School went from the lowest to the highest student satisfaction rating over a single year with a comprehensive program which involved assiduous communication with students, which included listening to their concerns and seeking to meet them.<sup>62</sup>

This having been said, much more could be made of unistats.com. It has obvious flaws, many of which could be ameliorated. And its functionality and usefulness could be substantially enhanced in myriad ways, to which we now turn.

## Improving existing models

This section explores some of the remaining shortcomings of unistats.com, many of which can be ameliorated, if not eliminated with better design and more collaborative use of technology. This process of critique naturally leads to suggestions which, if they were implemented in Australia's VET system and indeed more widely, could take Australia to the forefront of best practice.

- ✧ Asymmetric information is an important problem under the current system, which will only worsen with the deregulation of educational institutions. Those choosing courses often have sketchy knowledge about the quality of institutions and this may give an unfair and inefficient advantage to incumbents with a good reputation (whether deserved or otherwise).<sup>63</sup> In addition, 'peer effects' can mean that those institutions attracting the best students end up with better-measured outcomes, not because the teaching is better, but because students perform better in the presence of better, more stimulating peers. For this reason 'league tables' of educational institutions should reflect, not the absolute performance of their students, but some measure of the contribution or 'value added' by the institution in improving students' scores.<sup>64</sup>

- ✧ It would be possible to provide much more finely grained information if it were possible to interrogate the database. For instance, a user or researcher might want to know how a particular course was rated by students who performed highly or poorly and whether differences arose from the gender or ethnicity of students. We should work towards maximising access to all such data, subject only to restrictions that take account of important principles such as protection of privacy (these thoughts are expanded somewhat in the final section).

The appropriate principle is set out by Robinson et al. (2008, p.1). They argue that as a general rule, rather than ‘struggling ... to design sites that meet each end-user need’, governments ‘*should focus on creating a simple, reliable and publicly accessible infrastructure that “exposes” the underlying data*’ (emphasis in original).<sup>65</sup> Markets are often sufficiently shallow that it makes sense for governments to provide an interface with the data, although this should not compromise other potential suppliers’ access being built off the same architecture as proposed above. But where private providers of the data are viable, and they are likely to be in this case, there may be some sense in governments relinquishing their role of providing the interface to both save money and avoid ‘crowding out’ private provision.<sup>66</sup>

Some further issues arise from these considerations.

- ◆ If interfaces are to be provided privately—ratemyprofessors.com might provide an interface for the National Student Survey—it may be appropriate for such sites to be subject to some obligations; for instance, to ensure that advertising is subject to some code of practice. The issue here is that advertisers should not exert any influence on what information is presented, although contextual advertising *a la* Google ad words could be very valuable to users.
  - ◆ If it is necessary, there may be merit in publicly subsidising the development of certain display and analytic capabilities on websites. We discuss this further below.
- ✧ As with the Australian survey, a great deal that is of interest is not reported upon because there is a requirement that reported data achieve some basic level of statistical robustness. This raises a number of questions:
    - ◆ Firstly, the principle should be that more information is better than less. Where preferred statistical robustness has not been demonstrated, and, providing that a caveat is added, information of limited statistical significance should be made available.
    - ◆ Secondly, there is a variety of techniques for squeezing more information out of less. Some opinions are more equal than others. Highly favourable evaluations mean more from those who make them sparingly. Further, a comparison of a rater’s evaluations with the evaluations of others suggests how discriminating they are. The following diagrams illustrate an uncritically generous evaluator, whose preponderance of high ratings degrades the information value of any high ratings they provide, and a non-discriminating evaluator, whose ratings vary widely but not in a way that is consistent with others’ ratings. In each case we have strong reasons for suspecting the value of the rating.



**Figure 1 Uncritical and non-discriminating raters**



Source: Ho and Quinn (2008, p.283).<sup>67</sup>

By contrast Ho and Quinn (2008) argue that where something is highly rated by as few as two raters whose rating elsewhere shows them to be discriminating, this can nevertheless provide a relatively robust rating.

**Figure 2 Two discriminating raters**



Source: Ho and Quinn (2008, p.283).<sup>68</sup>

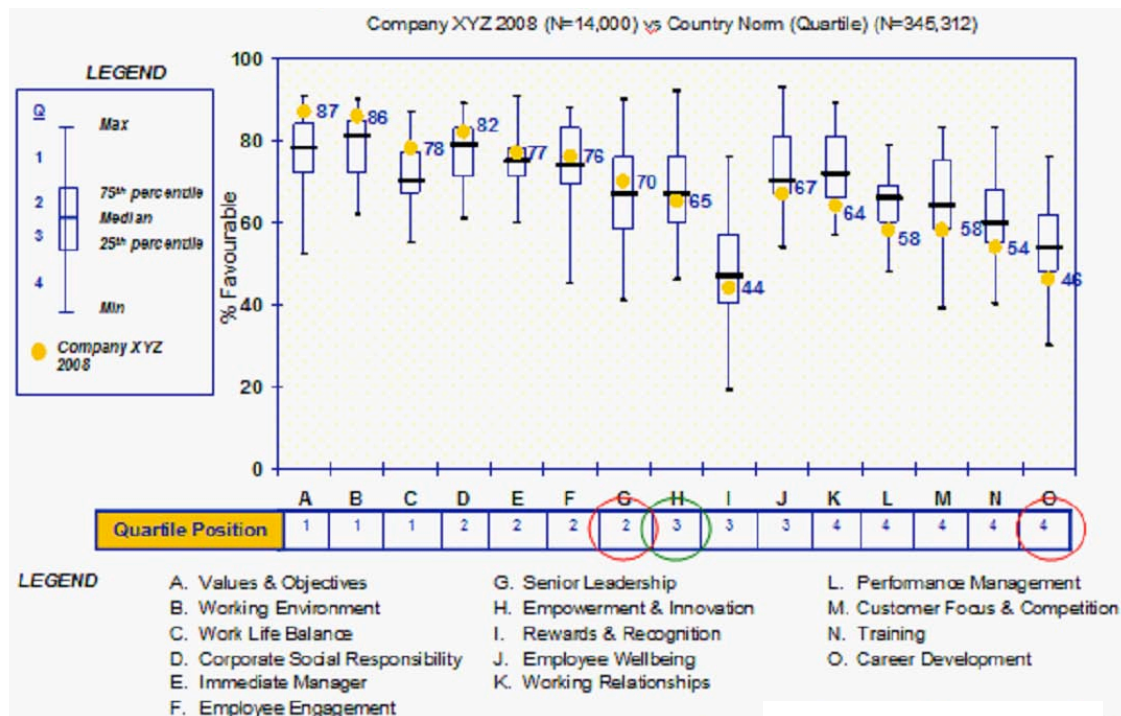
Such techniques evidently provide very powerful means for improving the statistical robustness of limited sample sizes.<sup>69</sup> An advantage that the official data systems like the National Student Survey and the Student Outcomes Survey have here is that, unlike the web-based private ratings providers like ratemyprofessors.com, they can reasonably vouchsafe the identity of all raters.<sup>70</sup>

✧ Statistical robustness is also improved by increased sample size. The United Kingdom's National Student Survey aims to be a full census of all students<sup>71</sup> and achieves a response rate of around 60%, with all institutions achieving a minimum response rate of 50%. We could replicate this in Australia, but of course without a change of methodology it would substantially increase costs. Current response rates to Australia's Student Outcomes Survey are 42.6% for the graduate survey and 33.2% for the Module Completed Component Survey.<sup>72</sup>

✧ There is a further way of improving response rates and this is to make responses compulsory. This seems reasonable, given that the information being collected is a public good. Furthermore, providing the information is opened to students for their own uses, it also seems fair.<sup>73</sup>

✧ Cognitive efficiency is also a crucial consideration. The *availability* of information is only the first step, with the next step being its clear communication to the user, which should include the user's ability to use, understand and manipulate it perceptively. For this reason much more effort should go into helping users appreciate various fine, and not so fine, points about using, interrogating and manipulating the information provided.<sup>74</sup> The issue of 'value added' has already been raised. At a more fundamental level, data should be provided in a way that helps users identify degrees of statistical significance and corresponding confidence intervals.<sup>75</sup> Further, some kinds of questions tend to attract lower ratings than others. Users of the data must know that if they are to use the data in a sensible way. In ratings of employee satisfaction, for instance, ratings of the adequacy of pay are typically unusually low compared with most other ratings of employee satisfaction. Thus for the data to be reported in a way that enables users to work out whether a firm has a *disproportionately* bad reputation for underpaying its employees would require a firm's raw score to be reported against industry averages (see figure 3).

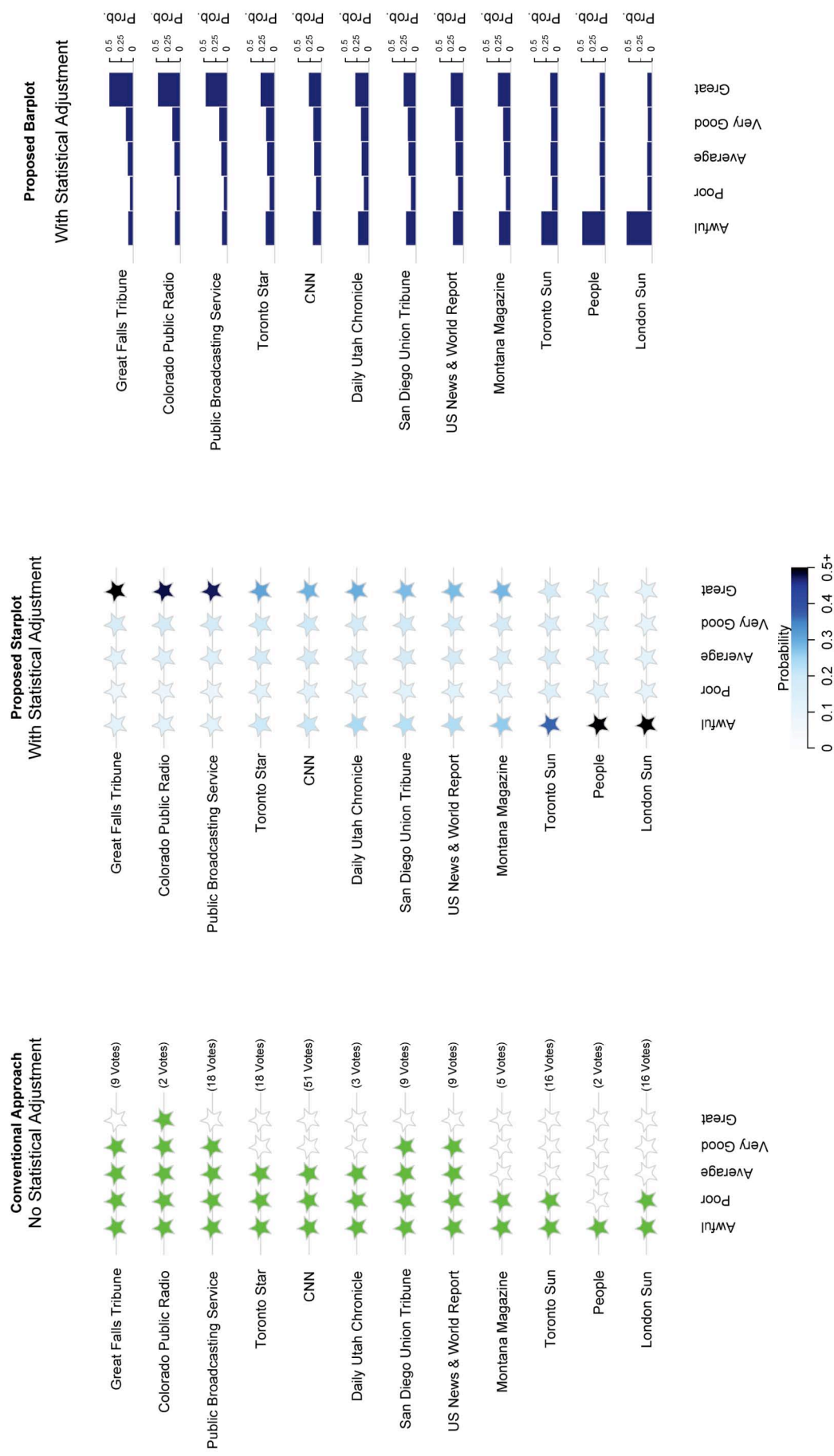
**Figure 3 Graphically illustrating the statistical context of various results**



Source: Personal correspondence with human resources firm.

How the data should be interpreted should at the very least be explained; however, it would be far better if users were able to access moderated scores according to professionally accepted and clearly documented methodologies. A variety of powerful ways to convey such information graphically is available and it is important to facilitate the use of such methods (see figures 3 and 4 for illustrative examples).

**Figure 4 Graphically presenting statistically richer data**



Source: Ho and Quinn (2008, p.282).

## Conclusion: Some speculations about the need to develop hybrid institutions, and the potential of Web 2.0

I conclude by returning to one of this paper's original themes. Hayek was concerned that collective economic institutions were insensible to the abundance of decentralised information throughout the economy, concluding that only markets could harness this information. His argument is appealing and has since been vindicated by experience. Yet on closer inspection, 'the market' itself—the mores, practices, laws and other institutions that support the existence of markets are a public good and in that sense a *collective* asset. Perhaps aware of the potential tension, Hayek, in his prosecution of his anti-socialist case, argued that the public goods which made markets possible—for instance, respect for private property and keeping a bargain—had spontaneously evolved as social norms over time within a culture. For him, the public good of the market was most efficacious and most secure if it was an instance of spontaneous order (what he called a *cosmos*), only subsequently legitimated and enforced by a centrally enforced mechanism (what he called a *taxis*).

Yet it was always the case that collectively—and centrally—imposed laws evolve alongside social experience, and not always simply legitimating what has already evolved. And since Hayek's heyday, more and more countries have refused to rely on the gradual evolution of spontaneous order, choosing instead to accelerate their transition to becoming sophisticated market economies by centrally designing and imposing the appropriate legal architecture.

This ecology between evolved and created social institutions has its counterpart in the hybrid institutions discussed at the outset of the paper, which comprise both private, individual and decentralised aspects, alongside collective and centralised aspects. The discussion in this paper suggests that the evaluations of individual students of their VET experience is a highly valuable resource. And yet, just as a market augments the productive power of its individual constituents by *inter alia* rationalising and standardising information, so the value of student evaluations can be harnessed much more effectively by deliberately designing and building the architecture of a system to do that, rather than relying on market forces alone, as ratemyprofessors.com must.

Having the capacity to vouchsafe the integrity of the identity of contributors, and indeed to compel their contributions, which is available only with the exercise of the collective power of the system itself, is a critical factor in a reconceptualisation of a database with the potential to provide accurate reputational information. The VET sector should move its current efforts with the Student Outcomes Survey in this direction, beginning with the publication of its results in an appropriately indexed and searchable form. But Web 2.0—the way in which the internet is now being used as a distributed IT platform between suppliers and users, producers and consumers and the way in which it can facilitate networking and collaboration between users—allows us to take things much further still. As Picci argues (2006, 2007a), Web 2.0 enables us to massively leverage 'word of mouth' information. And the process by which word of mouth is transformed into reputations is multidimensional, with communications not being between 'users' and a 'provider' or aggregator of information, but between just a few interested parties. This means that the traditional provider can also become a user. Usage patterns from unistats-style websites may just as readily enable institutions to learn more about their students and prospective students and what interests them.

Picci also distinguishes between 'ad hoc' and 'integrated' statistics. In Picci's terminology, the Student Outcomes Survey and indeed virtually all the statistics collected by the ABS are ad hoc, which is to say that they are discrete collections of data that have been specifically sought by the authorities. Against this Picci (2006, p.16) contrasts integrated statistics, which 'are produced as a view of the digital information already present within a computerized information system'. Ratemyprofessors.com provides integrated statistics in this sense, being continuously available both for receiving and sending data.

It seems a worthy goal—to work towards building a system in which the generation and dissemination of data on service quality is integrated with the delivery of post-secondary educational services. It is also likely that a properly integrated system would operate at low marginal costs (particularly where student input via the internet can be maximised, perhaps charges made for input supplied by other means). A system such as this would always be ‘on’ and available for occasional surveys during the teaching year—to give providers feedback on student responses to service delivery or for the consultation or involvement of students on other matters while courses were taking place.<sup>76</sup>

To further leverage word-of-mouth information, we can create a hierarchy of ‘opt ins’, which reflect different people’s preferences in the trade-offs between the protection of privacy and openness to others, as is now being pioneered on social networking sites. This is best illustrated with an example. Users could be allowed to adopt ‘avatars’ or internet identities chosen by them, which present them on the internet as a specific person, while preserving their anonymity to other users. However, this is a structured and not absolute or anarchic anonymity. To acquire an avatar they would undertake to communicate truthfully and in good faith. Their identities would be known to ‘the system’, so that their privileges could be modified or removed for misbehaviour and they could be pursued in the event of defamatory comments. They would also be warned that it may be possible for other users of the system to work out or speculate about their true identity.

For the sake of our example we have a student who is at the Mildura TAFE doing hospitality. He gives himself the avatar ‘Sunraysya’. When ‘Sunraysya’ contributes to discussion forums about the hospitality course at the Mildura TAFE, the system verifies that he is indeed qualified to comment; that is, that he is or has been a student in the relevant course.

Scientist Michael Neilson (2009) has commented on ‘the untapped creative potential existing in latent connections between scientists, and which could be released using suitable tools to activate the most valuable of those latent connections.’ Of course this is just an aspect of the greater value of human connectedness, something which is going through an epoch-making step change. Once this system of avatars and permissions is established, it becomes possible to facilitate the evolution of very socially, professionally and educationally useful networks of information and communication.

Information networks such as these do not currently exist, because the necessary ‘social networking’ technology is only just coming into common use on the internet, and because to date, statistical systems established by governments have typically imposed a ‘one size fits all’ set of privacy protections on users. Thus most statistical agencies have strict protocols for preventing the release of any information that might enable the identification of someone contributing data. Yet amongst those whose privacy is being protected, there exists a possibly substantial number who would be prepared to forego some privacy in return for others doing the same. Indeed, the way relationships typically develop—in our normal social lives or in cyberspace—is through a process of gradual and reciprocal revelation of information which remains private to others.

People could choose to establish ‘profiles’ either in their own name or in the name of an avatar and allow their profiles to be interrogated. They could elect to allow viewers of the profile to email them (either directly or via their avatar, which would still protect their anonymity); they could then respond as they wished—revealing their identity, responding still in the name of their avatar or ignoring the advance.

Such a system would facilitate the evolution of communities of interest and communities of common experience and would enable the deep mining of the database, where people might interrogate the system to identify whether a course or a teacher had been well regarded by ‘people like them’ in some specified respect(s), or search for those who had made the transition from one area of professional training to another. It would likewise enable teachers and course administrators to identify the strengths and weaknesses of an existing course and/or teacher, in terms of their appeal to different kinds of students, at a much greater level of detail than is possible today.

Of course this may remind readers of social networking sites like MySpace and Facebook. Facebook began in a tertiary institution—Harvard—with the initial goal of facilitating social, professional and pedagogical networking and communication. It has been built into a vast network with over 150 million users.<sup>77</sup> And Facebook now hosts applications, of precisely the kind—although I doubt yet of the scale—of what is being proposed here. It may well be that the most efficient and effective way to build the capability described here is not to build it on the *analogy* of Facebook, not to build it *like* Facebook, but to build it as an application *in* Facebook.

## References

- Becker, WE 2000, 'Teaching economics in the 21st century', *Journal of Economic Perspectives*, vol.14, no.1, pp.109–19.
- Bowling, NA 2008, 'Does the relationship between student ratings of course easiness and course quality vary across schools? The role of school academic rankings', *Assessment & Evaluation in Higher Education*, vol.33, no.4, August, 455–64.
- Centra, JA 2003, 'Will teachers receive higher student evaluations by giving higher grades and less course work?', *Research in Higher Education*, vol.44, pp.496–518.
- Coldarchi, T & Kornfield, I 2007, 'RateMyProfessors.com versus formal in-class student evaluations of teaching', *Practical Assessment, Research and Evaluation*, vol.12, no.6, pp.1–15, viewed 9 June 2009, <<http://pareonline.net/pdf/v12n6.pdf>>.
- Cruse, DB 1987, 'Student evaluations and the university professor: Caveat professor', *Higher Education*, vol.15, pp.723–37.
- Department of Education and Early Childhood Development, Victoria 2007, *Value-added measures for school improvement*, viewed 10 Jan 2008, <<http://www.eduweb.vic.gov.au/edulibrary/public/publ/research/publ/value-added-measures-report.pdf>>.
- Gillmore, GM & Greenwald, AG 1999, 'Using statistical adjustment to reduce biases in student ratings', *American Psychologist*, 54, pp.518–19.
- Ho, DE & Quinn, KM 2008, 'Improving the presentation and interpretation of online ratings data with model-based figures', *The American Statistician*, November 2008, vol.62, no.4, p.279.
- Hobson, SM & Talbot, DM 2001, 'Understanding student evaluations', *College Teaching*, vol.49, no.1, pp.26–31.
- Kay, J 2003, *The truth about markets: Why some nations are rich but most remain poor*, Penguin, London.
- Kindred, J & Mohammed, SN 2005, 'He will crush you like an academic Ninja! Exploring teacher ratings on Ratemyprofessors.com', *Journal of Computer Mediated Communication*, vol.10, no.3, article 9, viewed 3 January 2008, <<http://jcmc.indiana.edu/vol10/issue3/kindred.html>>.
- Knight, B & Cully, M 2007, 'A patchwork quilt: The development of national training statistics', in *Informing policy and practice in Australia's vocational education and training sector: Reflections and futures*, eds P Curtin & P Loveder, proceedings of the 25th anniversary forum of the National Centre for Vocational Education Research, NCVER, Adelaide, pp.24–31, viewed 8 January 2008, <<http://www.ncver.edu.au/publications/1830.html>>.
- McKeachie, WJ 1997, 'Student ratings: The validity of use', *American Psychologist*, vol.52, pp.1218–25.
- NCVER (National Centre for Vocational Educational Research) 2008, *Student Outcomes Survey 2008: Technical notes*, viewed 16 April 2009, <[www.ncver.edu.au/statistics/surveys/seuv05/seuv05\\_explanatory\\_notes.doc](http://www.ncver.edu.au/statistics/surveys/seuv05/seuv05_explanatory_notes.doc)>.
- Nielson, M 2009, *Connecting scientists to scientists*, viewed 29 January 2009, <<http://michaelnielsen.org/blog/?p=548>>.
- Picci, Lucio, 2006, 'Reputation-based governance of public works', *Rivista di Politica Economica*, no.I–II, pp.161–84, viewed 1 January 2009, <<http://didattica.spbo.unibo.it/picci/rebag-publicworks.pdf>>.
- 2007a, 'Reputation based governance', *First Monday*, vol.12, no.9, 3 September, viewed 1 January 2009, <<http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/viewArticle/2010/1885>>.
- 2007b, 'The reputational budget and its uses', presented at the First World Meeting of the Public Choice Society, Amsterdam, 29 March–1 April 1, viewed 2 January 2009, <<http://www.spbo.unibo.it/picci/rebag-reputationalbudget.pdf>>.
- Robinson, D, Yu, H, Zeller, W & Felten, EW (forthcoming), 'Government data and the invisible hand', in vol.11, *Yale Journal of Law and Technology* in Fall 2008, viewed 2 January 2008, <[http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=1138083](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1138083)>.
- Santiago, P, Tremblay, K, Basri, E & Arnal, E 2008, *Tertiary education for the knowledge society: Volume 1: Special features: Governance, funding, quality*, OECD, Paris.
- Stiglitz, JE (undated), 'Information', *The concise encyclopedia of economics*, viewed 16 April 2009, <<http://www.econlib.org/library/Enc/Information.html>>.
- United Kingdom, Department for Innovation, Universities and Skills 2003, *The future of higher education*, DIUS, London, viewed 17 January 2009, <<http://www.dcsf.gov.uk/hegateway/strategy/hestrategy/exec.shtml>>.

## Nicholas Gruen

- <sup>43</sup> Thanks to Damian Jeffree for valuable research assistance and Colin Alcock, Andrew Norton, Clive Kanés and David Kellam for their comments on an earlier draft.
- <sup>44</sup> The classic public good is defence from invasion of the realm because once the realm is defended it costs no more to defend all citizens (non-rivalry in consumption) and once one has protected one citizen one cannot exclude others from the protection (non-excludability).
- <sup>45</sup> A similar survey is done for the university sector in Australia, which is discussed briefly below.
- <sup>46</sup> McKeachie (1997, p.1219) argues that ‘student ratings are the single most valid source of data on teaching effectiveness.’ Hobson and Talbot’s 2001 review concluded that ‘well-developed student evaluations with adequate reliability and validity data may provide some of the best measures of teaching effectiveness’ (p.30). Centra (2003, pp.495–6) goes into more detail:
- No method of evaluating college teaching has been researched more than student evaluations, with well over 2,000 studies referenced in the ERIC system. The preponderance of these study results has been positive, concluding that the evaluations are: (a) reliable and stable; (b) valid when compared with student learning and other indicators of effective teaching; (c) multidimensional in terms of what they assess; (d) useful in improving teaching; and (e) only minimally affected by various course, teacher, or student characteristics that could bias results.
- <sup>47</sup> Note this information should not be ignored for many purposes, for instance, assessing course quality, but should be taken into account, for instance, when seeking to assess the quality of a teacher who has no control over class size.
- <sup>48</sup> In a ‘small year’ a random sample of around 82 000 is selected to be asked questions about their VET experience and their employment status. This provides statistically robust information at the level of states but not at the level of individual institutions or schools. Individual institutions can top up the sample from their student catchment at a marginal cost. In 2008, a small year, this increased the sample to about 108 000 students. In a large year, which produces institution-level reporting, the sample size is about 300 000 students.
- <sup>49</sup> There is some breakdown of results by ‘field of education’, although the information on student satisfaction and employment status does not seem to be adequately reported.
- <sup>50</sup> As Knight and Cully observe (2007, p.30), the current arrangement: ‘tends to favour the interests of the producers of vocational education and training (that is, the providers and the state training authorities) over the consumers (that is, individual students and employers) and the general public interest.’
- The protocols imposed on NCVER in its handling of information include the following:
- NCVER will not release information about an individual training provider without the written permission of that provider (or, where relevant, the appropriate state training authority). In addition, a number of state training authorities have requested that information about provider sectors within their jurisdiction (i.e. TAFE, community providers and private providers) not be published or released. As a result of these requests NCVER only publishes or releases sector-level information for those States and Territories that have agreed in writing to do so (viewed 8 January 2009, <<http://www.ncver.edu.au/aboutncver/statistics/protocols.html>>).
- This situation is similar to the one imposed upon the body managing a similar survey of university students in Australia (see below).
- <sup>51</sup> There will soon be some progress in the direction of online access to SOS information, with the March or April 2009 launch of SOS Online. This site, aimed at prospective students and modelled on GradsOnline (<[gradsonline.com.au](http://gradsonline.com.au)>), will feature information on employment outcomes and graduate salaries. Unfortunately, as per the restrictions on NCVER’s funding agreement, no information on a per-institution basis will be available, nor will any information relating to the graduates’ perceptions of teaching quality, effectively neutering the site as a driver of institutional competition (pers. comm. Susan Dawe and Mette Creaser [NCVER]).
- <sup>52</sup> <[Ratemyprofessors.com](http://Ratemyprofessors.com)> operates from five countries—but does not include Australia—from the US, Canada, England, Scotland and Wales.
- <sup>53</sup> Illustrating the principle that more information is usually better than less, it turns out that ‘hotness’ is correlated with higher-quality scores, offering a means for users to correct for the possibility of bias should they wish.

- <sup>54</sup> <<http://www.ratemyprofessors.com/About.jsp>>, viewed 17 January 2009.
- <sup>55</sup> For instance, giving professors a specific ‘channel’ to voice their responses to student feedback (or putative student feedback!) and with rejection of defamatory and obviously poorly motivated ratings and comments.
- <sup>56</sup> See <<http://www.thestudentsurvey.com/>> for further detail.
- <sup>57</sup> <<http://www.unistats.com/faq.do#s2q1>>. There are plans to collect and publish information at course level in future. The closest thing Australia has to the NSS is the Course Experience Questionnaire organised by Graduate Careers Australia. The teaching quality score for each area of study at each institution is given as average, worse, or better and is published in the *Good universities guide*. Graduate Careers Australia is a not-for-profit organisation run by representatives from the universities themselves and government. The course-level teacher-quality rating is currently not published online, although the data are made available to the universities. The level of compression of the data in the *Good universities guide* makes it effectively useless as a basis of competition between university course offerings. In fact, the universities are expressly prevented from using the Course Experience Questionnaire data to compete against each other by a clause in the contract with Graduate Careers Australia. By comparison, the English NSS is commissioned by the Higher Education Funding Council for England, which is the funding body for all higher education in England, and the study has been designed to facilitate informed competition between the institutions. Nevertheless, it is possible that the data gathered by the Course Experience Questionnaire could be co-opted as the basis for a Unistats-style site (with thanks to Andrew Norton for information on the Course Experience Questionnaire).
- <sup>58</sup> <<http://www.timeshighereducation.co.uk/story.asp?storyCode=310073&sectioncode=26>>, viewed 17 January 2009.
- <sup>59</sup> See for instance <<http://mhslearning.edublogs.org/2008/09/26/do-you-teach-no-i-work-at-the-university/>>, viewed 10 January 2009.
- <sup>60</sup> See the case studies at <[www.heacademy.ac.uk/assets/York/documents/ourwork/research/surveys/nss/nss\\_case\\_studies\\_nov07\\_v5.doc](http://www.heacademy.ac.uk/assets/York/documents/ourwork/research/surveys/nss/nss_case_studies_nov07_v5.doc)>.
- <sup>61</sup> See for instance De Montfort University, <[http://www.ipsos-mori.com/\\_assets/nss/pdf/de%20montfort.pdf](http://www.ipsos-mori.com/_assets/nss/pdf/de%20montfort.pdf)>. Part of its plan, communicated in some Powerpoint slides, involves ‘develop[ing] own branding about the survey’ and making sure students are ‘back [and] settled in’ (I presume this means when the survey is administered.)
- <sup>62</sup> As reported on a blog post by Kevin O’Brien, Professor of Orthodontics:
- So what did we do?*
- Head of School, Dr Grey took hands on approach with high student visibility.
  - Concentrated on being positive and pushing the good things.
  - Improved communication. The students received the Head of School weekly update.
  - Started a BLOG.
  - We had an excellent student year rep with whom we worked. She knew that she could approach any of us and we would listen to her and help.
  - Staff/student meeting. We made a point of listening to the students and acting on their concerns. There had been a tendency to ‘brush issues’ aside in recent years.
  - Let the students know how we had responded to their concerns and requests. Extra lectures, extra clinics. There was transparency.
  - One member of staff/tutor met with a group of 10–12 students for an hour every week. This was the personalised contact between the group and the tutor. We tended to let the group set the agenda. If the group had nothing to fill the time we had prepared material to go through with them. That member of staff made sure that the session occupied the hour.
  - The tutors met 3 times, with their students individually to go through their log book/portfolios and give feedback, advice etc.
  - The tutors made sure that the students realised that they were receiving feedback. We changed the heading on the sheets to read—feedback.
  - Each group of students had a rep that would speak on behalf of the group, for the group. These reps also met with the year coordinator.
  - The tutors also met with the year coordinator to [deal with] feedback issues. These were quickly addressed and actions reported back to the students.



- The year coordinator was supported by a full time administrator/secretary. They both had open door policies. The students appreciated this.
- The Head of School had 1 to 1 emails—he invited questions from the students, to which he responded personally.
- More staff student interactions/events.

Basically, we listened to the students, responded to requests and let them know what we had done or not done (in a few cases) <<http://mhslearning.edublogs.org/2008/09/26/do-you-teach-no-i-work-at-the-university/>>.

<sup>63</sup> Andrew Norton has commented, I think rightly, that this is more of a problem in the universities than it is in VET.

<sup>64</sup> See, for instance, Department of Education and Early Childhood Development (2007). Note that the problem of ‘established’ institutions enjoying self-reinforcing reputational advantages by virtue of their incumbency seems much less pronounced a problem in the VET sector than it could be in the university sector.

It may not be appropriate to develop such corrections for such things in student satisfaction with courses (although it is worth trying to further disaggregate results to provide some profile of the kinds of students who rate themselves as satisfied and unsatisfied with courses). However, where league tables reflect the numbers of students finding employment after completing a course, one would have thought it important to provide users of unistats.com with some means of inferring the extent to which outcomes are influenced by the quality of the institution as opposed to the quality of its students. Not doing so is likely to perpetuate unfair and inefficient advantages of incumbent institutions.

<sup>65</sup> They go on:

Private actors, either nonprofit or commercial, are better suited to deliver government information to citizens and can constantly create and reshape the tools individuals use to find and leverage public data. The best way to ensure that the government allows private parties to compete on equal terms in the provision of government data is to *require that federal websites themselves use the same open systems for accessing the underlying data as they make available to the public at large* (emphasis in original).

<sup>66</sup> Note however that there is a further problem. If privately run sites are generating data, for instance, in discussion forums or through the additional provision of information by users, this may fragment the available information, unless the various competing sites are nevertheless coordinated in some way, perhaps taking data feeds from one another.

<sup>67</sup> As Ho and Quinn (2008, p.283) explain the diagrams and what they are illustrating.

[Figure 1] plots the observed ratings for two raters of *U.S. News*. Each panel represents all ratings submitted by two raters on the *y*-axis (randomly jittered for visibility) and the [posterior mean of the] latent content quality as estimated from the IRT model on the *x*-axis. The first panel shows that *U.S. News* was rated by a non-critical user, who rated more than two-thirds of all outlets as ‘great’ (i.e., better than *U.S. News*). If anything, from this user’s rating of ‘very good’ we learn that *U.S. News* may be worse than the majority of outlets rated. The second panel plots a user who is largely non-discriminating, failing to distinguish high and low content outlets in any systematic way compared to the majority of Mondo users. Intuitively, we learn little from such users, as a rating of ‘very good’ does not distinguish the outlet meaningfully. Little information is conveyed ...

<sup>68</sup> All ratings submitted by the two raters, both of whom rated *Colorado Public Radio* in row 2 of figure 2. Each panel depicts data from a single rater. Each circle in each panel represents a news outlet rated by the rater in question, randomly jittered for visibility within each of the five rating categories. The filled blue circles represent the rating of *Colorado Public Radio*. This figure illustrates how, even with only two ratings, the (posterior) probability that this news outlet would be rated ‘great’ is quite high. Note that both raters rated many outlets and both raters are very good at distinguishing low- and high-quality news outlets.

<sup>69</sup> Note: to the extent that this technique is used, it presumes an objective standard of rating. Accordingly, the statistical approach would need to be re-specified and would lose some power, to the extent that we sought to interrogate the data for more than one perspective; for instance, a high-performing student’s perspective versus a lower-performing student’s perspective.

<sup>70</sup> On the other hand, ratemyprofessors.com has a strong incentive to use such techniques, which can disrupt and raise the cost of ‘strategic’ or manipulative posting. If ratemyprofessors.com downgraded

those raters who show erratic ratings, those seeking to make their ratings count would have an incentive to try to match them to their expectations of others' ratings.

- <sup>71</sup> Student instructions are offered which imply that completion of the survey is in the first instance compulsory; for example, by outlining procedures for 'opting out', but there is no discussion of any sanctions other than follow-up reminders.
- <sup>72</sup> See <<http://www.hefce.ac.uk/learning/nss/>> and <<http://www.hefce.ac.uk/news/hefce/2008/nss.htm>>, viewed 16 January 2009 for data on the NSS and NCVER (2008, p.7) for SOS data.
- <sup>73</sup> The ABS has powers of compulsion in recognition of this argument, although, if such a process were sensibly adopted, the cumbersome procedures that the ABS is required to go through would be avoided (the individual provision of specific directions to provide information followed by court enforcement). A more promising approach would be analogous to that taken in enforcing compulsory voting. Someone required to provide information would be required on pain of some fine (and/or as a condition of the granting of their qualification) to register to provide the information but retaining some residual right, having gone to that inconvenience, to not provide the information. We do not actually compel voting, we compel the inconvenience of voting while allowing people to vote informally once they have submitted to the inconvenience. Similar principles should apply here. Although it is not truly compulsory, the UK student survey incorporates similar principles to those being proposed here. The student survey website <[studentsurvey.com](http://studentsurvey.com)> makes no reference to any ultimate sanctions for non-completion of the survey, but it implies that the survey is compulsory, by providing advice on 'opting out' of the survey in the following terms:

We need to be sure of the identity of the student opting-out and ask for the same identification information for opting-out of the survey as we do for responding to the survey. To opt-out of the survey click here. You will be able to opt-out of specific stages of the NSS (e.g. the online survey), or from all stages.

It then goes on to pose another question, 'I have already completed the survey and I recently received a reminder. Do I have to complete it again?'. Its answer explains what students to do—implying, without saying so explicitly—that responses are compulsory <<http://www.thestudentsurvey.com/faqs.asp>>, questions 11 and 14, viewed on 17 January 2009.

A further issue here is that, if the provision of student evaluations is to be made compulsory, the survey and the methods available to students for completing it should not be onerous. One way this could be assured would be to adopt the British practice, which is to involve student associations in the governance of the system.

- <sup>74</sup> Currently NCVER does allow approved users access to a confidentialised form of the raw data through confidentialised unit record files (CURFs). In addition, VOCSTATS <<http://www.ncver.edu.au/resources/vocstats/intro.html>> allows sophisticated users to construct their own tables extracted from the various NCVER collections, including the SOS via a web interface (pers. comm. Susan Dawe).
- <sup>75</sup> NCVER publishes confidence levels for all SOS statistics published online. VOCSTATS does not generate confidence levels but does have a starring system to indicate data quality.
- <sup>76</sup> It might also facilitate more comprehensive consideration of more students' interests in the construction of course timetables. Pucci (2006, p.17) also argues that integrated statistics are more robust to political exigencies than ad hoc statistics.

While the degree of institutionalization of *ad hoc* statistics greatly varies, all integrated statistics tend to be highly institutionalized. Once they have been produced for the first time, their future provision only depends on the continuing existence of the information system, and on the maintenance of the software producing the necessary views of the data. In order to discontinue the production of integrated statistics a clear opposition is needed, and not just a lack of support. To put it differently, *ad hoc* statistics requiring funding need an explicit decision to be continued. On the other hand, integrated statistics, once the continued functioning of the underlying information system is guaranteed, need an explicit decision in order to be interrupted.

- <sup>77</sup> <<http://blog.facebook.com/blog.php?post=46881667130>>, viewed 18 January 2009.