A matter of degrees

Rowan Myron-Wilson

Peter K. Smith

Rowan Myron-Wilson *and* **Peter K. Smith** *look at the rising number of 'good' degrees in psychology and at how departments vary.*

ONE index of teaching quality in British universities has been the proportion of 'good' degrees awarded — firsts and upper seconds. Obtaining a good degree has also been important for many students seeking postgraduate qualifications and awards. Despite the increase in numbers in higher education, and some decrease in funding per student, the proportion of good degrees awarded has risen steadily since the 1970s.

The Higher Education Quality Council (1996) has detailed these changes in eight subjects. Over all eight there has been a modest rise in first class degrees since the mid-1980s, but a more rapid rise in the proportion of upper seconds. This proportion was below 35 per cent into the early 1980s, but then rose steadily to around 45 per cent in 1993 (HEQC, 1996, Figure 3.1). The proportion of both lower second and third class degrees fell correspondingly during this period. In addition, within a discipline, there is considerable variation between institutions.

Modelling variability

Chapman (1993, 1994, 1996), who was the author of the HEQC report, had earlier analysed the situation regarding geography, and had developed a model for variability in degree outcome. This identifies 'inputs' (student characteristics and quality), institutional influences (resources, programmes, degree norms), system influences (external examiners, codes of practice) and attributes of departments (resources, teaching structure, assessment procedures) as influencing the 'output' of degree class distributions.

A portion of the variability in degree classes can indeed be ascribed to student intake quality, as defined by A-level scores. Nevertheless, taking account of student intake, it is then possible to assign each institution a 'value-added' or 'value-subtracted' position, depending on the proportion of good degrees it is giving in that discipline. Positions of departments with respect to value added or subtracted could be explained by variations in other parts of Chapman's model — for example, in teaching quality and resources, or in practices relating to degree class norms and criteria.

Some earlier analyses have been made of the situation in psychology. Connolly and Smith (1986) analysed degree class distributions in the 'old' universities, based on data from a partial sample for 1977–80 and a virtually complete sample for 1981–84. In the latter period, the variation in proportion of good degrees across departments was from 22 to 73 per cent. The correlation with mean A-level score was only 0.17, and there were similarly modest correlations with other indices of teaching quality and resources, and with research productivity.

Smith (1990) produced a further analysis covering the period 1985–88. The range in proportion of good degrees was still very marked: from 29 to 77 per cent in the 'old' universities, and from 17 to 49 per cent in the public sector. Correlations with mean A-level scores were only 0.24 or 0.25 within each sector, but rose to 0.55 when combined, no doubt due to the greater variance obtained. Analyses of measures of teaching quality and resources, and research excellence, both by correlation and stepwise regression, still yielded only modest and not very consistent findings.

Nevertheless, the proportion of good degrees appeared to be rising in psychology departments over this period. The distribution for the 'old' universities was higher in 1985–88 than in 1981–84; there was also a significant increase in the whole sample during the 1985–88 period. A worrying possibility, regarded by Smith (1990, p.151) as 'not yet disproved', was that both the variation between departments and the changes over time might reflect norms and criteria for degree class allocation, more than student intake characteristics or aspects of teaching quality and resources.

In the light of the renewed interest at the national level (HEQC, 1996), we decided to re-examine the data for psychology departments eight years later, and with the binary divide between universities and the former polytechnics removed. With the co-operation of the Association of Heads of Psychology Departments, we contacted psychology degree-awarding institutions in England, Wales, Scotland, Northern Ireland and the Republic of Ireland, with a questionnaire slightly modified from that used in 1990.

Gathering the data

The degree questionnaire began with the breakdown of degree classes for each year (1989 through to 1996). The questionnaire then asked for mean A-level

scores (or equivalent) for entrants through central admissions services (like the present Universities and Colleges Admissions Service), the proportion of mature students, selection processes, staff–student ratios, tutorial arrangements, assessment methods (including the extent of double internal and blind marking, measured on 3-point scales), and the perceived role and actions of the external examiner.

The questionnaire was distributed in the UK and the Republic of Ireland to all heads of departments awarding degrees in which psychology was the single or main component (eligible for the Society's Graduate Basis for Registration (GBR)). The questionnaire was sent to 92 departments; 85 were returned. Of these, nine were not included in the analysis: four because they were not eligible for GBR, and five because their degree programme, while eligible for GBR, had not yet reached first graduation. This left a maximum of 76 responses for analysis (fewer for some analyses as not all responses were complete).

Much of the analysis focuses on the percentage of 'good' degrees given, defined as firsts and upper second class degrees. In line with the time periods used in previous studies (Connolly and Smith, 1986; Smith, 1990), we present analyses separately for our two four-year periods, 1989–1992 being labelled T1, and 1993–1996 being labelled T2.

Influences on performance

Student characteristics

Few departments returned usable data on mean A-level grades (as point scores): from only 14 for 1989 to 36 for 1994. The mean A-level point scores rose steadily from 19.74 in 1989 to 21.25 in 1995, but dropped again in 1996 to a mean score of 20.39.

	1989–1992				1993–1996			
	1st	**2:1**	**2:2**	**3rd**	**1st**	**2:1**	**2:2**	**3rd**
more	21	12	5	0	53	13	7	7
no change	77	84	91	98	42	79	89	87
fewer	2	4	4	2	5	8	3	6

Table 1: Percentage of institutions whose external examiners advised them on the whole to give more, the same number, or fewer of each degree class

The percentage of mature students taken ranged from 1 to 60 per cent (with one institution an exception at 100 per cent). The mean intake of mature students rose from 20.91 per cent (T1) to 23.75 per cent (T2).

Teaching characteristics

The mean staff–student ratio was 1:10.74 in T1 and 1:9.74 in T2. The size of tutorials varied greatly over departments, with a range from one to 25, and a mean tutorial size of 10.98; most institutions were in the range five to 15. Most departments reported that their tutorials were given by both staff and postgraduates (31) or by staff alone (27). In only two institutions were tutorials given by postgraduates alone.

Most institutions reported giving tutorials in all years of the degree (45); eight departments gave tutorials only in the first year, and nine only in the first and second years. A considerable proportion of departments reported having an optional structure within their degree programme (81 per cent in T1, 89 per cent in T2).

Assessment characteristics

In the first four-year period, 51 departments reported using two or more internal examiners (13 answered no, and 12 did not respond). This did not significantly change in T2, when 57 departments reported using two or more internal examiners (14 answered no, and five did not respond). However, a marked change was noted in the extent of blind marking. In T1 only 30 departments out of 64 reported using blind marking; this figure rose significantly in T2 to 51 out of 71 departments.

Degree criteria

Departments were asked whether explicit criteria for assigning marks to individual items of work were available to internal examiners. Just over half the departments (35 out of 61) in T1 did make such criteria available. This figure rose considerably to 51 (out of 70) departments in T2. Many departments (48 out of 69) also reported that they made these criteria available to students.

External examiners

There was a great variety of views on the role of the external examiner. The most common roles were seen as those of comparing standards across institutions (23), of giving advice on standards (16) and of adjudicating difficult cases (13); while 10 departments said they viewed their external examiners as having a variety of roles (usually including the ones above). Three saw validation of classification as the main role, and three saw it as giving critical feedback.

Departments frequently reported that external examiners had made comments and given advice on degree marking and awarding: 29 in T1, and 41 in T2. The way in which these comments were reported to have influenced degree classifications is shown in Table 1.

The advice of external examiners has predominantly been not to change the existing degree distribution but, where they have been perceived to recommend change, was usually to recommend more firsts. This was especially so in T2, however there was no significant difference when compared with T1. They have also more often tended to recommend increasing, rather than decreasing, the proportion of upper seconds, while being more balanced about lower seconds and thirds.

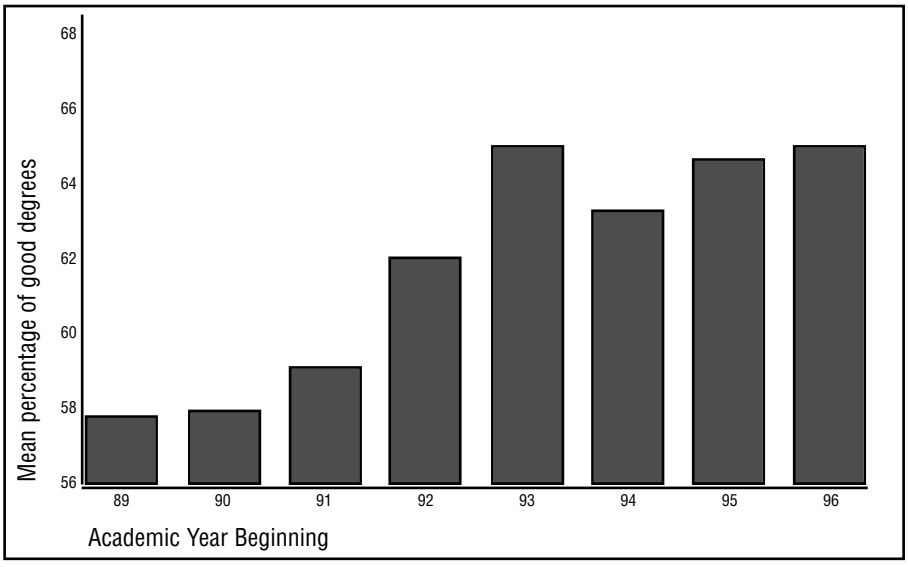

Figure 1: Percentage of good degrees for all departments

Variations in classification

Changes over time

There were significant changes over time in the distribution of degree classes over

the two four-year periods. The proportion of firsts rose from 5 per cent in T1 to 7 per cent in T2; the proportion of upper seconds also rose, from 54 to 58 per cent. By contrast, the proportion of lower seconds dropped from 37 to 32 per cent, and of thirds by 4 per cent to 3 per cent. The mean percentage of good degrees rose from 59 to 65 per cent between the two periods. Figure 1 demonstrates the general increase over the two periods in the percentage of good degrees awarded.

There was some variation in the proportion of good degrees depending on the status and area of the institution (see Table 2). However, in all cases, there is an increase from T1 to T2.

In both periods, old universities awarded a significantly higher percentage of good degrees than new universities. In T1, the old universities gave an average of 62 per cent of good degrees, compared with 51 per cent awarded by newer universities. In T2, this gap had widened with old universities awarding 69 per cent, while new universities awarded 55 per cent.

Variations between departments

There was wide variation across different institutions in the proportion of good degrees awarded. For T1, the range was 33.1 to 83.7 per cent, and in T2 29.7 to 89.5 per cent. This range of some 50 per cent between the extremes is much the same as in the earlier studies, though again the distribution has shifted upwards a few percentage points.

The consistency of individual departments over time is demonstrated by the high correlation of percentages of good degrees in T1 and T2 (r = .72). The range of variation is also illustrated in Figure 2, which averages the T1 and T2 data.

Degree class correlates

The percentage of good degrees at each institution (averaged over each of the two four-year periods) was correlated with the following variables:

- mean A-level score
- percentage of mature students
- staff–student ratio
- mean tutorial size
- amount of optional choice in course structure
- extent of double internal marking
- extent of blind marking
- perceived direction of external examiner's advice
- Research Assessment Exercise (RAE) ratings from each period (1989 for T1, 1992 for T2).

1989–1992 (T1)

The only significant correlate with the percentage of good degrees was the (1989) RAE rating (r = .45), which in turn correlated quite strongly with mean A-level score (r = .51). Nevertheless, the correlation of percentage of good degrees with A-level scores of entrants was small (r = .25) and non-significant. Note, however, that few institutions returned A-level score data for this period.

	1989–1992	1993–1996
Old universities (43)	61.9	69.2
New universities (16)	51.3	54.6
England & Wales (48)	59.7	66.0
Scotland (8)	60.4	62.2
Northern Ireland (2)	52.7	65.4
Republic of Ireland (4)	54.6	60.9

Table 2: Percentage of good degrees across type of institution and location of institution

The mean A-level score also correlated moderately with the amount of optional course structure (r = .50) and the amount of double internal marking (r = .45).

These results may indicate that students with better A-level grades prefer departments with a stronger research base and with better resourced and flexible courses. However, it does little to explain the variance among departments in proportion of good degrees.

1993–1996 (T2)

The percentage of good degrees again correlated significantly with departments' (1992) RAE ratings (r = .52). It also correlated negatively with the percentage of mature students (r = –.25) and with tutorial size (r = –.29), indicating that the smaller the tutorials and the fewer the mature students in the group, the higher the percentage of good degrees.

The RAE rating again correlated significantly with A-level score (r = .70) and with staff–student ratio (r = .52), and negatively with the percentage of mature students (r = –.32) and tutorial size (r = –.29). However, the correlation of good degrees with A-level scores of entrants was again small (r = .11) and non-significant.

To examine which variables were independently predictive of the percentage of good degrees, we carried out a stepwise regression. The dependent variable was taken as percentage of good degrees, the independent variables as tutorial size, staff–student ratio, RAE rating, amount of blind marking, double marking and optional structure in the course, and the perceived direction of the external examiners' comments (scored on a 3-point scale).

For T1, the regression was not significantly different from zero; none of the variables was independently predictive of the proportion of good degrees. For T2, it was significantly different from zero. However, only one of the independent variables contributed significantly to prediction of good degrees: RAE rating. Tutorial size showed a trend towards significance .

Consistent trends

The first main finding from this analysis is that the proportion of good degrees has continued to rise over the period 1989 to 1996. This has been a fairly steady trend over this eight-year period (Figure 1), continuing the trend in 1977–80, 1981–84 and 1985–88 (Connolly & Smith, 1986; Smith, 1990). It is consistent for both old and new universities (Table 2). It is also consistent with the overall trend for eight other degree subjects, analysed in the HEQC (1996) report.

The second main finding is that the wide variation between psychology departments in the proportion of good degrees awarded, noted in the earlier surveys, is still just as marked. There also

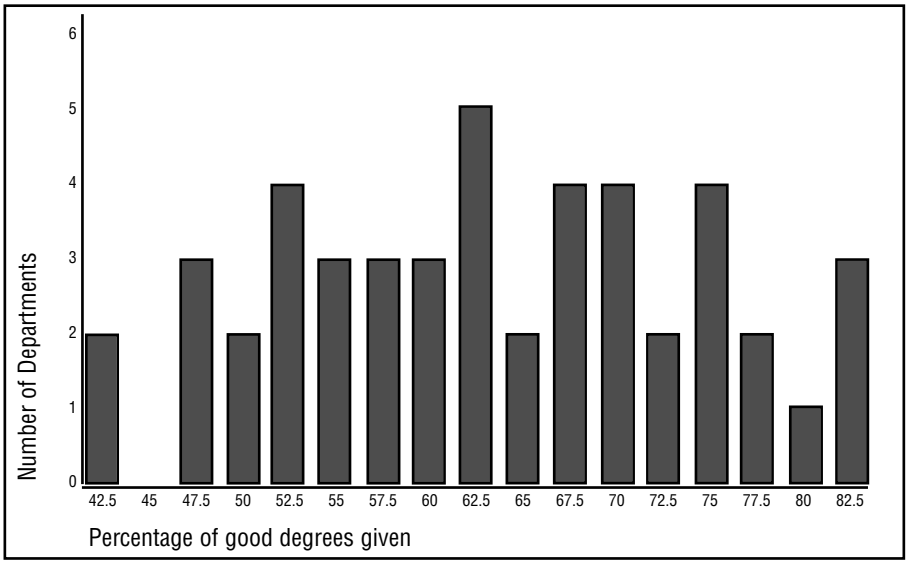

Figure 2: Department variation in percentage of good degrees (1989–1996)

appears to be high consistency over time in these inter-departmental variations.

What can explain these two main findings?

Why do institutions vary?

It is apparent that variability between departments has persisted for some two decades now. It has proved difficult to find correlates, but the consistent correlate in this survey, significant also in the T2 regression analysis, is with research productivity as indexed by the RAE. This correlate appears to be increasing in significance with time (it is more significant in T2 than in T1, and both these correlations are more significant than that of r = .32 reported in Smith (1990) for the 1985–88 period).

There are some slight, but not very consistent, indications that aspects of teaching quality such as options, double marking and tutorial size, may be influences — but only tutorial size appears as even marginally significant in one of the regression analyses.

A tentative explanation of this pattern of results might be as follows. The successive research assessment exercises have become both more valid, and more widely known to prospective psychology students. They have also become more strongly linked to financial benefit for more highly-rated departments. As one widely known external indicator, students may use RAE ratings as a basis for applications, and good students are more likely to go to highly-rated departments (as confirmed by the positive correlations obtained between these two variables).

These more highly-rated departments thus get better qualified students, and may also be able to afford smaller tutorial groups and a better staff–student ratio (see pattern of correlations in T2). These probably have some impact on the proportion of good degrees, although as we have seen, their influence may not be large or consistent.

Why more 'good' degrees?

The other issue is of the continuing increase in the proportion of good degrees. It is now modal for students to get an upper second class degree, whereas 20 years ago it was modal to get a lower second class degree. Given the fall in real funding per student over this period, it seems unlikely that this change is solely explicable in terms of improvements in teaching quality, difficult as this is to measure (a conclusion also reached in HEQC, 1996).

An optimistic view is that some of this improvement is due to teaching

quality, organisation and resources. Better textbooks are being produced, lecturers have more training, and students are becoming more vocal and assertive consumers, which together with quality assurance procedures puts pressure on departments to teach well. Teaching syllabuses and examinations may be better linked, and modular structures may make exam questions more predictable. Semesterisation may help students pace their learning, receiving more feedback and working more consistently through the academic year.

An alternative explanation is that much of the increase comes from a realisation in institutions that the proportion of good degrees is a public indicator of success. This indicator has been used, for example, by HEQC in teaching assessments, and sometimes by research councils. Departments with a relatively low level of good degrees may come under pressure to improve this.

The role of external examiners, which has often been seen as one of encouraging departments to be more generous with first class degrees, has over the last decade included some encouragement to increase upper second class degrees as well (Table 1). Thus, the external examiner system has probably provided some legitimisation for this national increase and helped diffuse an expectation of a modal upper second class degree.

Overall, it would appear that there are some weak links between the percentage of good degrees and aspects of teaching quality, such as tutorial size and quality of students as indexed by A-level scores. However, these may be mediated by RAE scores, with, paradoxically, a measure of research productivity emerging as the most robust indicator of an apparent measure of teaching success — the proportion of good degrees awarded.

Nevertheless, many institutions may have an internal culture of awarding a certain proportion of good degrees, contributing much to the inter-departmental variation obtained. These internal cultures of marking have, nationwide, been affected by the incentives to increase the proportion of good degrees awarded, and this increase has largely been validated by the external examiner system, at least as perceived by the recipients of their advice. In terms of Chapman's (1993, 1994) model, we may have a fairly complex interaction between 'input', institutional influences, system influences and attributes of departments, in producing 'output'.

Achieving consistency

Perhaps the most relevant recommendation made by the Dearing Report (National Committee of Inquiry into Higher Education, 1997) regarding standards is to amend the remit of the Quality Assurance Agency (formerly the HEQC) to include verifying standards and maintaining the 'qualification framework' — seen by Dearing as embodying nationally applicable criteria for a comprehensive and comprehensible system of awards in higher education.

The emphasis the Dearing Report places upon teaching standards and the recommended creation of an Institute for Learning and Teaching in Higher Education will perhaps also go some way towards reducing idiosyncratic variation. If, as recommended, all full-time academic staff are required to achieve at least associate membership of the Institute, this may provide an opportunity for consistency across institutions which currently appears to be lacking. This may be especially true if the Institute encompasses criteria and standards within its remit.

References

Chapman, K. (1993). Degree results in geography 1973–1990: Students, teachers and standards. *Area*, **25**, 117–126.

Chapman, K. (1994). Variations of degree results in geography in United Kingdom universities 1973–90: Preliminary results and policy implications. *Studies in Higher Education,* **19**, 89–102.

Chapman, K. (1996). Entry qualifications, degree results and value-added in UK universities. *Oxford Review of Education,* **22**, 251–264.

Connolly, K.J. & Smith, P.K. (1986). What makes a good degree: Variations between different departments. *Bulletin of the British Psychological Society*, **39**, 48–51.

Higher Education Quality Council (1996). *Inter-institutional Variability of Degree Results: An Analysis in Selected Subjects.* London: Higher Education Quality Council.

National Committee of Inquiry into Higher Education (1997). *Higher Education in the Learning Society.* London: The Stationery Office.

Smith, P.K. (1990). The distribution of psychology degree classes in the UK. *The Psychologist,* **4**, 147–152.

Rowan Myron-Wilson is at the Department of Psychology, University of Hull, Hull HU6 7RX. Tel: 01482 465472; e-mail: r.r.myron-wilson@psy.hull.ac.uk.

Professor Peter K. Smith is at the Department of Psychology, Goldsmiths College, University of London, New Cross, London SE14 6NW. Tel: 0171 919 7870; e-mail: pss01@gold.ac.uk.

A full account of this research can be obtained from the authors.