

Performance of British University Psychology Departments as Measured by Number of Publications in BPS Journals

ANDREW M. COLMAN, SONYA P. GRANT, AND EMMA HENDERSON

University of Leicester

The aim of this investigation was to gauge the comparative research performance of university departments of psychology in Britain. The performance indicator was the number of publications per departmental staff member in the journals of the British Psychological Society (BPS) during the seven-year period 1983–1989. The contents of these journals were thoroughly searched, and articles written by members of university psychology departments in Britain were counted. The number of publications of each psychology department was divided by the size of the department in the corresponding year and the resulting annual figures were summed to determine the number of publications in BPS journals per departmental staff member over the seven-year period. These research performance figures correlated significantly with an earlier measure of departmental research performance based on publications in the seven European psychological journals with the highest citations per published article and with recent performance indicators based on other criteria

The funding of British universities since the mid-1980s has been increasingly influenced by ratings of departmental research performance. Although selectivity is becoming commonplace, there is little agreement about how departmental research performance should be measured. The two major research selectivity exercises carried out by the University Grants Committee in the 1980s were severely criticized on logical and empirical grounds, and some constructive suggestions were made (Bentham, 1987; Gillett, 1987; Gillett & Aitkenhead, 1987; Johnes, 1988).

In discussions about performance indicators, there often has been confusion between departmental performance on the one hand and departmental productivity, or output, on the other (Gillett, 1989). Performance, properly understood, is a measure of cost-effectiveness inasmuch as departments with identical research outputs, however output is measured, obviously differ in performance if their sizes and disposable resources differ. Performance indicators are thus measures that relate output (productivity) to input (resources). The most straightforward output measures are simple publication counts, which are crude measures of the quantity of research output, or citation counts, which are assumed to take both quantity and quality into account. The most obvious input factors are the number of departmental staff members, the size of equipment grants and other recurrent incomes, and the amount of research grant income. Departmental research grant income was apparently used as a performance indicator in the

1985–86 and 1989 University Grants Committee research selectivity exercises, and it continues to be used a such, although it is not really a performance indicator at all. It is obviously an input measure which takes no account whatever of the quantity or quality of research output. Research grant income has, in fact, been found to bear virtually no relation to output measures such as publication or citation counts (Bentham, 1987; Gillett, 1987). To use research grant income as a performance indicator rather than an input variable is to confuse means and ends. A suitable performance indicator should measure research output per unit of input. The performance indicator used in the investigation described below is the number of publications in British Psychological Society (BPS) journals per staff member from 1983 to 1989.

A simple publication count (not restricted to any specific set of journals) provides the most direct and straightforward measure of research output, but there are problems associated with this measure. Most importantly, it weights all publications equally without taking into account differences in length and quality. Publication counts that are restricted to refereed journal articles iron out some of the gross differences (between books and articles, for example) but still treat long and important articles in leading journals on a par with brief and trivial contributions in obscure, low-quality journals.

Citation counting, which is thought by some researchers to overcome the problem of quality control, is based on the assumption that the number of times a publication is cited by subsequent authors in journal articles is—or can be used as—an index of its quality. Citation counting has been widely used to measure research output for over twenty years, especially in the United States (Endler, Rushton, & Roediger, 1978; Garfield, 1979, 1987; Rushton, 1984; Rushton, 1989). A departmental citation count is usually defined in practice as follows: the number of citations, in articles published during a specified period in journals covered by the *Social Science Citation Index* or one of the other citation indexes, of publications by members of a department. Citation counts are thought by many to provide better measures of output than do simple publication counts, because the frequency with which a publication is cited can be taken as an indirect measure of its impact in the scientific community and therefore of its quality. Citation counts, like publication counts, are certainly objective and quantitative indices of departmental research performance, and provided that input variables such as departmental size are properly taken into account, they can be used to derive performance indicators that are to be preferred to subjective, impressionistic ratings. But there are problems with this method of measuring departmental performance. Chapman (1989) listed 25 shortcomings, biases, deficiencies, and limitations which, taken together, amount to a major indictment of citation counts as performance indicators.

A more recent method of dealing with the problem of quality control associated with simple publication counting is by limiting publication counts to papers in a restricted set of high-quality journals. This methodology was first used by Cox and Catt (1977) and has more recently been repeated by Howard, Cole, and Maxwell (1987) in a study of research performance in the United States based on publications in the 13 journals of the American Psychological Association (APA) and by Colman, Garner, and Jolly

(1992) in a study of the research performance of United Kingdom psychology departments based on the seven European psychological journals with the highest mean citations per published article. Although some of the articles published in leading journals are inevitably better than others, the assumption underlying this methodology is that all of the articles are likely to be above a certain high threshold of quality to have been accepted for publication in the best journals. This new methodology avoids many of the problems of citation counting identified by Chapman (1989). Thus, problems arising from the fact that the *Social Science Citation Index* and the *Science Citation Index* cover slightly different branches of psychology fall away. Trivial publications such as letters, conference abstracts, and book reviews, which are given full weight in citation counts, can be easily excluded; the quality of the periodicals carrying the publications, which is inadequately controlled in citation counting, is guaranteed by the initial choice; all authors, rather than only first-named authors, as in citation counting, can be given due credit; the overwhelming influence in citation counts of individual "stars" is avoided, because the variance in numbers of articles is vastly less than the variance in numbers of citations; mis-counting of citations because of slightly different versions of an author's name or initials is avoided; the well known bias of citation counts against newcomers, which is due to the lag between the appearance of an article and the relevant citations, is eliminated; tactical manipulation of citations counts by self-citation or mutual citation conspiracies ceases to be a worry; the problem of spuriously inflated citation counts associated with papers reporting new tests, statistical procedures, and other methods or recipes falls away; and the fact that a paper can accumulate many citations from commentators who disapprove of its contents does not distort the results as it does with citation counting. There are problems that remain, of course, and other problems that are peculiar to the new methodology; some of them will be discussed later. What seems clear, however, is that the new methodology avoids many of the problems and has certain advantages over both citation counting and simple publication counting.

METHOD

This research was essentially a replication of the American studies of Cox and Catt (1977) and Howard, Cole, & Maxwell (1987) which assessed the research performance of institutions in the United States by counting papers in the 13 journals of the American Psychological Association. To assess the research performance of university departments of psychology in Britain, we focused on the seven journals of the British Psychological Society (BPS): the *British Journal of Psychology*, the *British Journal of Social Psychology*, the *British Journal of Clinical Psychology*, the *British Journal of Mathematical and Statistical Psychology*, the *British Journal of Developmental Psychology*, the *Journal of Occupational Psychology*, and the *British Journal of Medical Psychology*. Each of these journals was thoroughly searched for the years 1983 to 1989 inclusive for articles written by members of British university psychology departments. The reason for starting in 1983 rather than 1980 was that the *British Journal of Develop-*

mental Psychology and the *British Journal of Social Psychology* split at the beginning of that year. Credit was assigned to authors, and hence to departments, according to the following formula, borrowed from Howard, Cole, and Maxwell (1987), which is designed to give appropriate weight to the number of authors and their relative rank order:

$$\text{credit}_i = (1.5^{n-i}) / (\sum 1.5^{i-1}),$$

where i is the specified author's ordinal position, n is the total number of authors, and the summation is over i from $i = 1$ to n . For a single-authored paper, the score was therefore 1; for a double-authored paper, the first author scored 0.6 and the second author 0.4; for a triple-authored paper the three authors scored 0.47, 0.32, and 0.21 respectively, and so on. No credits were given for ephemera such as letters, book reviews, software reviews, and editorial notes.

The 42 psychology departments included in this study are listed in Table 1. Data for London RHBNC (Royal Holloway and Bedford New College) include publications by members of London University's Bedford College before it merged with Royal Holloway College (which had no psychology department). Publications from UWIST (University of Wales Institute of Science and Technology) were assigned to University College Cardiff with which UWIST merged in 1989. In the case of the University of Sussex, the departments of developmental, social, and experimental psychology were not distinguished.

The size of each department for each of the years from 1983 to 1989 inclusive was determined from information supplied in the corresponding *Commonwealth Universities Yearbook*. Data for the University of London Institute of Psychiatry, which is not listed in the *Yearbook*, were obtained directly from the department itself. Credits were once again assigned roughly proportionally: full-time staff members scored one full unit, and part-time staff members and those holding shared appointments in two departments scored one-half a unit. Staff members on temporary and fixed-term appointments were counted in the same way as tenured staff for the period of their appointments. Research assistants and fellows, visiting and honorary members of staff, demonstrators, and emeritus professors were not counted. There were occasional ambiguities about departmental sizes in the information contained in the *Yearbook*, but the figures eventually obtained are almost certainly more accurate than the estimates, based mainly on figures in the 1985 APA *Directory* and assumed fixed for 1976–1985, that were used by Howard, Cole, and Maxwell (1987) in their American study.

On a year-by-year basis, the author credit score for each department was divided by the corresponding staff size figure to yield an annual mean performance score, and these annual means were then summed to yield an overall performance score for the full seven-year period.

RESULTS

The final performance scores, in descending rank order, are shown in Table 1 together with performance scores calculated in a similar way by Colman, Garner, and Jolly (1992) using articles in the seven European (EC) psychology journals with the highest

TABLE 1
Performance (BPS Journals and European Community Journals),
Publications, and Citations of British Psychology Departments

<i>Institution</i>	Performance					
	BPS Journals		EC Journals ^a		Publications ^b	Citations ^b
	<i>Mean</i>	<i>Rank</i>	<i>Mean</i>	<i>Rank</i>	<i>Mean</i>	<i>Mean</i>
Sheffield	2.34	1	0.75	14	2.3	18.7
Kent	2.14	2	0.00	41.5	3.8	11.8
Oxford	1.53	3	3.33	1	2.1	45.5
Exeter	1.45	4	1.83	3	2.3	17.9
Leicester	1.25	5	1.53	6	2.4	12.2
London: Birkbeck	1.10	6.5	2.05	2	1.5	21.7
London: Inst. Psych.	1.10	6.5	0.38	33	-.-	-.-
Lancaster	0.97	8	1.71	4	2.6	15.0
London: UCL	0.91	9	0.92	12	2.3	15.0
Birmingham	0.84	10	0.61	18.5	2.8	20.9
Dundee	0.81	11	1.31	9	2.6	7.1
York	0.77	12	0.64	16	1.8	13.3
Warwick	0.76	13	0.39	31.5	0.5	4.5
Cambridge	0.74	14	0.74	15	2.6	46.5
Bristol	0.72	15	0.50	22	1.0	13.0
Aberdeen	0.70	16	0.53	20	0.4	5.1
Surrey	0.66	17.5	0.05	40	1.3	5.7
Sussex	0.66	17.5	1.64	5	2.9	14.8
Liverpool	0.64	19	0.33	34.5	0.8	8.8
Durham	0.62	20	1.25	10	0.7	10.6
London: RHBNC	0.60	21	1.33	8	1.7	4.4
Nottingham	0.59	22	1.39	7	2.0	11.2

^a Data for 1980-1989 from Colman, Garner, & Jolly (1992), Table 1.

^b Data for 1985 only from Rushton (1989), Table 1.

TABLE 1 (cont'd)

Wales: Bangor	0.56	23	0.61	18.5	1.2	10.4
Keele	0.52	24	0.62	17	1.4	7.5
Glasgow	0.45	25	0.22	36	0.5	3.3
Strathclyde	0.43	26	0.40	30	0.5	10.1
Stirling	0.40	27	0.42	27	2.1	10.1
London: LSE	0.39	28.5	0.43	25.5	0.8	5.3
Wales: Cardiff	0.39	28.5	0.51	21	1.8	10.7
Manchester	0.36	30	0.46	23.5	0.7	5.1
Reading	0.35	31	1.15	11	1.6	6.4
St Andrew's	0.34	32	0.39	31.5	0.6	11.1
Queen's Belfast	0.31	33	0.19	38	1.3	2.8
Ulster	0.29	34	0.43	25.5	1.5	4.2
Aston	0.24	35	0.46	23.5	1.2	4.0
Leeds	0.20	36	0.41	28.5	1.2	11.3
Edinburgh	0.17	37.5	0.33	34.5	1.3	19.6
Wales: Swansea	0.17	37.5	0.79	13	1.2	8.7
Southampton	0.15	40	0.41	28.5	0.5	2.7
Newcastle	0.14	41	0.13	39	0.2	5.2
Hull	0.09	42	0.21	37	0.7	2.8
Bradford	0.07	43	0.00	41.5	1.8	9.4
Mean	0.67		0.76		1.52	11.57
Standard Deviation	0.48		0.66		0.82	9.32

mean citations per published article, three of which were BPS journals. Also shown in Table 1 are one-year mean publication and mean citation counts, both based on data contained in the 1985 *Social Science Citation Index* (Rushton, 1989).

Pearson's product-moment correlation coefficients were calculated between the performance scores based on the seven BPS journals used in this study (column 1 in Table 1) and the various other performance estimates shown in the table. The correlation between BPS journal performance (1983-1989) and performance in leading European psychology journals (1980-1989) is $r(40) = 0.43$ ($p < 0.01$). The Spearman rank correlation between the ranks on the same two variables is $r_s = 0.51$ ($p < 0.01$). The correlation between BPS journal performance (1983-1989) and publications in all journals covered by the *Social Science Citation Index* in 1985 is $r(39) = 0.62$ ($p < 0.001$). Finally, the correlation between BPS journal performance (1983-1989) and citations in journals covered by the *Social Science Citation Index* in 1985 is $r(39) = 0.46$ ($p < 0.01$). None of the differences between these correlations is statistically significant.

DISCUSSION AND CONCLUSIONS

The departmental research performance scores reported in this paper were defined as the number of publications per departmental staff member in the seven journals of the British Psychological Society over the seven-year period 1983 to 1989. The results show a considerable range of performance among university departments of psychology in Britain, with the universities of Sheffield, Kent, and Oxford in the top-ranking positions. The performance scores correlate significantly, in the range $r = 0.43$ to $r = 0.62$, with three recent measures of departmental research performance based on different productivity measures.

The correlations would have been considerably higher if it were not for a small number of apparent anomalies. The most striking of these is the University of Kent, which ranked second for performance in BPS journals but equal bottom for performance in the seven European psychology journals with the highest mean citations per published article (Colman, Garner, & Jolly, 1992). It is not difficult to explain this apparent anomaly. The Kent department is a department of social psychology, and virtually all of its output in BPS journals between 1983 and 1989 was accounted for by articles in the *British Journal of Social Psychology*. On the other hand, none of the leading European psychology journals used in the previous study (the *British Journal of Psychology*, *Acta Psychologica*, the *British Journal of Educational Psychology*, the *British Journal of Medical Psychology*, *Ergonomics*, the *Quarterly Journal of Experimental Psychology*, and the *British Journal of Mathematical and Statistical Psychology*) specializes in social psychology, and that explains why Kent ranked near the top on the measure used in this investigation and very much lower on the measure used in the earlier investigation. Similarly, the Sheffield department, which has special strength (through a research unit attached to it) in occupational psychology, was the highest ranked of all 42 departments according to the BPS list, but ranked only 14th according to the list of leading European journals, largely because of its considerable output in the *BPS Journal of Occupational Psychology* and the absence of any journal specializ-

ing in occupational psychology on the European journals list. The University of London Institute of Psychiatry achieved a much higher rank according to the BPS set than the European set largely because of its considerable output in the BPS *British Journal of Clinical Psychology* and the absence of any clinical psychology journal in the European set.

For reasons set out in the introduction to this article, the performance measure used in the present investigation—the number of papers per departmental staff member published between 1983 and 1989 inclusive in the seven journals of the British Psychological Society—has certain advantages over simple publication counts and citation counts. The results of applying such a measure depend to some extent, of course, on the particular set of journals used. The set of leading European psychology journals used by Colman, Garner, and Jolly (1992) appears to have disadvantaged a few departments with particular strengths in areas not adequately covered by those journals. The seven journals of the British Psychological Society used in this investigation, on the other hand, covered most important areas of pure and applied psychology, and this seems to have attenuated the problem. It is worth pointing out that the rank correlation between the two sets of results was none the less quite high ($r_s = 0.51, p < 0.01$), which shows that rather similar ranks were assigned despite the very different journal sets used in the two studies. It is also important to realize that any performance indicator that might be devised is bound to be somewhat arbitrary and conjectural, which means that no two performance indicators are likely to rank all departments identically.

Although the performance indicator used in this study avoids many of the objections that have been raised against simple publication and citation counts, there are grounds on which it can be criticized. First, it takes no account of the fact that some of the articles in the chosen journals are longer and better than others. However, although the peer review process which decides the fate of manuscripts for publication is notoriously unreliable (Harnad, 1982), when attention is restricted to journals such as those used in this study, all of which have high rejection rates, a degree of quality control is ensured because few if any genuinely weak papers are likely to have survived the peer review process and to be included in the results. A second objection is that the performance measure used in this study completely ignores publications other than journal articles, with the result that it gives no credit for research output in the form of edited books, chapters in edited books, and monographs. It is impossible, however, to take everything into account without running into serious methodological problems. The results reported in this paper show, in any event, that performance estimates based on restricted journal sets correlate quite highly with each other and with more comprehensive indices of performance, including citation counts. American researchers (Howard, 1983; Howard, Maxwell, Berra, & Sternitzke, 1985; Maxwell & Howard, 1986) have reported evidence that confirms this conclusion.

It is worth pointing out in conclusion that a university department is a multivariate organizational structure in terms of both inputs and outputs, and that any publication-based performance indicator provides only a partial indication of overall performance, since it entirely ignores teaching and other departmental activities. To that extent it can

be misleading if it is overinterpreted, and the same criticism applies to citation-based performance indicators.

NOTES

Date of acceptance for publication: July 22, 1991. Address for correspondence: Dr. A.M. Colman, Department of Psychology, University of Leicester, Leicester LE1 7RH, England.

REFERENCES

- Bentham, G. (1987). An evaluation of the UGC's ratings of the research of British university geography departments. *Area*, 19, 147-154.
- Chapman, A.J. (1989). Assessing research: Citation-count shortcomings. *The Psychologist: Bulletin of the British Psychological Society*, 8, 336-344.
- Colman, A.M., Garner, A.B., & Jolly, S. (1992). Research Performance of United Kingdom Psychology Departments. *Studies in Higher Education* 17, 97-104.
- Cox, W.M. & Catt, V. (1977). Productivity ratings of graduate programs in psychology based upon publication in the journals of the American Psychological Association. *American Psychologist*, 32, 793-813.
- Endler, N.S., Rushton, J.P., & Roediger, H.L. III. (1978). Productivity and scholarly impact (citations) of British, Canadian, and US departments of psychology (1975). *American Psychologist*, 33, 1064-1082.
- Garfield, E. (1979). Is citation analysis a legitimate analysis tool? *Scientometrics*, 4, 359-375.
- Garfield, E. (1987). Mapping the world of science: Is citation analysis a legitimate evaluation tool? In D.N. Jackson & J.P. Rushton (Eds), *Scientific Excellence: Origins and Assessment*. Beverley Hills, CA: Sage.
- Gillett, R. (1987). Serious anomalies in the UGC comparative evaluation of the research performance of psychology departments. *Bulletin of the British Psychological Society*, 40, 42-49.
- Gillett, R. (1989). Research performance indicators based on peer review: A critical analysis. *Higher Education Quarterly*, 43, 20-38.
- Gillett, R. & Aitkenhead, M. (1987). Rank injustice in academic research. *Nature*, 327, 381-382.
- Harnad, S. (Ed.). (1982). *Peer Commentary on Peer Review: A Case Study in Scientific Quality Control*. Cambridge: Cambridge University Press.
- Howard, G.S. (1983). Research productivity in counseling psychology: An update and generalization study. *Journal of Counseling Psychology*, 30, 600-602.
- Howard, G.S., Cole, D.A., & Maxwell, S.E. (1987). Research productivity in psychology based on publication in the journals of the American Psychological Association. *American Psychologist*, 42, 975-986.
- Howard, G.S., Maxwell, S.E., Berra, S.M., & Sternitzke, M.E. (1985). Institutional research productivity in industrial/organizational psychology. *Journal of Applied Psychology*, 70, 233-236.
- Johnes, G. (1988). Research performance indicators in the university sector. *Higher Education Quarterly*, 42, 54-71.
- Maxwell, S.E. & Howard, G.S. (1986). Research productivity in quantitative psychology. *Educational and Psychological Measurement*, 46, 845-852.
- Rushton, J.P. (1984). Evaluating research eminence in psychology: The construct validity of citation counts. *Bulletin of the British Psychological Society*, 37, 33-36.
- Rushton, J.P. (1989). A ten-year scientometric revisit of British psychology departments. *The Psychologist: Bulletin of the British Psychological Society*, 2, 64-68.