

Advancing Public Accountability? The Social Services 'Star' Ratings

Tony Cutler and Barbara Waine

An important development in the UK public sector is the rating and classification of service providers. The authors examine a significant exercise of this kind: the 'star' ratings awarded to local authority social services departments in England. The article explains why it is difficult to accept the Government's claim that these star ratings provide a significant advance in public accountability.

The attempt to measure the performance of public sector service providers is not a new phenomenon (Cutler and Waine, 1997) but, as Pollitt and Bouckaert (2000) argue, it has become more significant in three dimensions:

- Performance measurement is more 'extended' in terms of the range of services and the dimensions of performance covered.
- It is more 'intense', as performance data has moved from being background information to being used in the setting of standards, and is frequently linked to explicit rewards and penalties.
- Performance information has moved from being for internal organizational purposes to being for 'public consumption' (*ibid.*; see also Broadbent, 2003), thus in discussing the rationale for performance measurement a report from the National Audit Office claimed that 'good quality information...enables people to participate in government' and therefore has a role in 'empowering citizens' and 'improving accountability' (NAO, 2001).

This article focuses on the production of performance assessment for public consumption and examines the issues raised by this practice through a consideration of a recent important exercise—the social services performance 'star ratings'.

The Production of the Star Ratings

Star ratings for social services departments in England were first published in May 2002. Of the 150 local authorities covered, eight were awarded the highest rank of three stars, 50 were classified as two stars, 82 as one star and 10 were designated zero stars (SSI, 2002a). In a reclassification exercise in November 2002, the ratings were amended, with 11 authorities awarded three stars, 52 receiving two stars, 75

given one star and 12 zero stars (SSI, 2002b). The ratings were produced by the Department of Health's Social Services Inspectorate (SSI) using a methodology which combined quantitative and qualitative elements (see SSI, 2002c and d). The quantitative element was provided through the 50 performance indicators of the Social Services Performance Assessment Framework (SSPAF) and, in particular, a subset of 11 indicators which were defined as Key Performance Indicators (KPIs) for the star ratings assessment (DoH, 2002b).

The use of KPIs was intended to give a definite role to quantitative measures in the assessment process. Thus the SSI argued that their use would ensure that 'performance indicators had sufficient weight in the star ratings system'. In addition, as a common set of indicators was being used, local councils would be 'treated in the same way' (SSI, 2002a). With respect to designation of which indicators were to be used as KPIs, selection was said to be based on:

- 'Importance' (the relation of the indicator to the Government's priorities).
- 'Ease of interpretation'—the indicator did not 'require further contextual data'.
- Data that was 'believed to be reliable'.
- Data that was 'attributable to social services', i.e. the performance was 'largely due to the performance of social services rather than other factors or agencies' (SSI, 2002c).

The qualitative dimension of the methodology was provided by the SSI's inspection reports and by Joint Review teams managed by the SSI and the Audit Commission (for a discussion of Joint Reviews, see Humphrey, 2002). By reviewing the evidence, and attempting to ensure consistency by using a set of descriptors of good and poor performance, the inspectors make their

Tony Cutler is Senior Lecturer in Management and the Programme Director of the Health Services Management MSc, Royal Holloway, University of London.

Barbara Waine is Honorary Research Fellow, Department of Health and Social Care, Royal Holloway, University of London.

judgement on services for children and adults in relation to current performance and future prospects (SSI, 2002d). With respect to current performance, the assessment focuses on whether 'services are serving people well', and the performance categories are 'no', 'some', 'most' and 'yes'. With respect to expected future performance, the judgement is of prospects for service improvement and these can be classified as 'poor', 'uncertain', 'promising' and 'excellent'. This performance matrix is converted into a final 'star' rating using a predetermined set of rules. Three principles underlie these rules:

- Current performance is rated more important than prospects for improvement.
- Services for children and adults are given equal weighting.
- An overall failure in either adult services or services for children means that the department is not 'serving people well' and this will result in zero stars (SSI, 2002d).

Star Ratings: The Rationale

The social services star ratings involved a complex process and this raises the question of the rationale for the exercise. Two justifications are given by SSI. The first is to give information to key stakeholders—service users, the general public, taxpayers and central government—regarding the performance of individual social services departments (SSI, 2002a and d). The second is to provide a means of differentiating good and failing performers, so that what is seen as the appropriate management regime is applied. The best-performing departments (with three stars) will have an 'increasing level of freedom in the way they use centrally-provided grant funds' and will experience 'a lighter touch programme of inspection and monitoring, and reduced requirements for planning information'. The poorest performing departments (zero stars): 'can expect increased levels of intervention, and will be subject to more rigorous and frequent monitoring' (SSI, 2002a). This approach parallels the one adopted in the National Health Service (NHS) where star ratings for acute trusts were first issued in September 2001.

The publication of the ratings is indicative of the importance attributed to their use for public accountability, and the ratings are claimed to represent an advance over previous practice. They are distinguished from the previous regime of performance assessment in social services in three ways:

- Information is said to be presented in a more accessible way, which is seen as essential if the performance assessment is to reflect, in particular, the needs of service users (Platt, 2002).
- The measures will change from being partial summaries of performance to 'more rounded assessments' and, as such, will use a wide range of evidence—performance indicators, inspection reports, plans and evidence from monitoring.
- They involve classification of departmental performance which will be 'a new part of the assessment process, involving both judgements about current performance but also [an assessment of] prospects for improvement' (Platt, 2002; SSI, 2002c and d).

Improving Public Accountability?

How far, then, do the social services star ratings represent an advance in public accountability? Commenting on the star ratings, the Chief Inspector of the SSI, Denise Platt (2002), has stated that they 'make a simple statement about the performance of a fiendishly complex service but they serve an important purpose—letting the public know how well their local authorities perform'. However, this statement raises an issue fundamental to problems of performance assessment. Given that judgements on service provision are 'fiendishly complex', does the attempt to render them in the form of a 'simple statement' lead to a presentation which is misleading?

The star system used was conducive to 'good' headlines. Thus the Chief Inspector of the SSI commented that the media had commented favourably that performance ratings had been found 'easy to understand'. However, if the star ratings are to advance public accountability, they need to pass a series of tests. An understanding of the value of the star system requires a knowledge of how the ratings were derived. To properly understand the methodology and allocation of ratings, it is necessary to consult at least four SSI publications (SSI, 2002a–d). A paradox of the aspiration to more 'rounded' assessment is that it would also be necessary to consult the relevant Joint Review and SSI reports for individual local authorities.

Even with access to all these sources there is still a problem with interpretation, i.e. the ability to evaluate the different types of evidence, in particular the performance indicators and inspection data used. What appears to be implied in the SSI reports on the star ratings process is that, since the process of producing

the star is so systematic, it is possible to move seamlessly from evidence to assessment to classification. This is in contrast to a substantial academic literature which points to the difficulty of making rigorous connections between quantitative performance measures and the objectives of public services (see Boyne, 1997; Cutler and Waine, 1997; Pollitt and Bouckaert, 2000; in relation to SSPAF, see Plank, 2000; Miller, 2002). It has frequently been observed that performance indicator sets either do not, or only inadequately, reflect service performance (Boyne, 1997). The response to this perceived problem has been to supplement quantitative measures with qualitative assessments of service provision. As was indicated above, this approach is reflected in the social services star ratings which have been designed to use *both* performance indicators and evidence from inspections.

The inspection reports are provided by the SSI and Joint Reviews: both of these bodies are charged with evaluating the quality of social services provision and do this via a methodology that is predominantly qualitative, for example visits, interviews with users and carers, examination of policy documents, tracking individual referrals through the system. While this approach has merit, it does present a number of problems and a central issue is consistency. The teams for each SSI and Joint Review vary, since the quantity of inspections (three SSI inspections in a five-year period and a Joint Review every five years) make it impossible to have the same people carrying out each review (a similar point was made by Hood *et al.*, 1998 in respect of Ofsted inspections). Thus it is essential to have a method of ensuring consistency of judgement and criteria to guide inspectors. Even then, it cannot be assumed that the adoption of a systematic set of criteria or protocols will necessarily result in consistent judgements. Consider, for example, the criteria used by the Joint Review teams in making judgements of how well social services are serving the local population. An essential one is that authorities must show that 'services enable people to live independent lives in as safe an environment as is reasonable given their personal circumstances' (Joint Review, 2001). This would appear to raise issues of how key terms such as 'independent lives', 'as safe an environment as is reasonable' are to be defined. Either these are treated as matters for the discretion of the Joint Review team or there is an attempt to add sub-descriptors (for example a definition of 'independent'), which arguably are ineffective in imposing constraints on such

judgements—a feature which has been discussed in the case of Ofsted inspections (Gilroy and Wilcox, 1997).

At the time that the star ratings were published in May 2002, only 73% of all social services departments in England had had a Joint Review, and the reviews of all local authorities were not due to be completed until the end of 2003 (Joint Review, 2001). Bexley, a three-star authority, had not, at the time when the star ratings were published, had a review. This raises doubts regarding the claim that departments are being assessed by a common yardstick. Some have been evaluated using indicators, SSI reports and joint reviews, others only indicators and SSI reports. Similar problems have arisen in the NHS. Clinical governance reports from the Commission for Health Improvement (CHI) have affected the ratings of NHS acute trusts (Smith, 2002). However, in 2001/02, of 46 three-star acute trusts, 30 had not had a CHI review published between September 2001 and July 2002. Only CHI reports published in this period were taken into account in assessing ratings (DoH, 2002a).

Reliance on inspection data for a limited period is therefore likely to result in patchy coverage. There are also problems in extending such time periods, since this can mean that inspections do not refer to a common time period. Because of the review cycle, some social services departments had had their reviews as early as 1997 and could have either improved or deteriorated in the intervening period. Haringey, for example, was reviewed in 1999 and received, on the whole, a favourable review, which concluded that 'the residents of Haringey and the users of social services are generally well served' and that 'there is no reason why the Authority should be anything but optimistic about its future' (Joint Review, 1999). In December 2000, Haringey was put on special measures and in May 2002 was zero rated, being viewed as serving only some adults and no children well and having poor prospects for improvement for both groups (SSI, 2002a).

A further problem relates to assessment of performance changes. The NAO (2001) suggests that part of the empowering role of performance measurement is that it allows for citizen pressure for 'continuous improvement'. The corollary is that it is possible to see if improvement has occurred and this, in turn, requires consistency in performance measures. As only one social services star ratings exercise has taken place, it is not yet possible to see whether changes in measures will be a source

of confusion. However, the precedents in health are not encouraging. The second star ratings exercise covering acute trusts changed the way in which one key target was defined and extended the range of supplementary 'broader' indicators from 12 to 28 measures (DoH, 2002a). Thus, there would appear to be difficulties in accepting the claim that the star ratings provide a significant advance in public accountability.

Conclusion: Revisiting Public and Managerial Accountability

The social services star ratings suggest that the two dimensions of public and managerial accountability are inextricably and positively linked. In this approach, published data explain to the public how the ratings were reached and these then trigger the 'appropriate' managerial regime, ranging from markedly reduced supervision (three stars) to much tighter supervision and, at the limit, a change in management (zero stars).

However, there is arguably a fundamental problem in the postulated relationship between public and managerial accountability that undermines this approach. The social services star ratings involve an ambiguity in assessment which is familiar to academic commentators on performance evaluation in public services generally. However, the ratings are presented as if a 'simple' account can be achieved providing a negative impact on managerial accountability. This raises the question as to what drives this quixotic quest for simplicity and certainty. Arguably, the star ratings, in Pollitt and Bouckaert's terms, exhibit 'intensity'; there are sticks and carrots. This suggests a need to legitimize the eventual ratings as clear and unproblematic. Yet this is inconsistent with the imprecision which has been revealed by an examination of the assessment process. Complexity in performance assessment would appear to be inescapable—trying to reduce this to a simple grid as a means of producing managerial accountability seems a questionable way of advancing public accountability. ■

References

Boyne, G. (1997), Comparing the performance of local authorities: An evaluation of the Audit Commission indicators. *Local Government Studies*, 23, 4, pp. 17–43.

Broadbent, J. (2003), Comprehensive Performance Assessment: The crock of gold at the end of the performance rainbow? *Public*

Money & Management, 23, 1, pp. 5–7.

Cutler, T. and Waine, B. (1997), *Managing the Welfare State: Text and Sourcebook* (Berg, Oxford).

Department of Health (2002a), *NHS Performance Ratings: Acute Trusts, Specialist Trusts, Ambulance Trusts, Mental Health Trusts 2001/02* (London).

Department of Health (2002b), *Social Services Performance Assessment Framework Indicators 2001–2002* (London).

Gilroy, P. and Wilcox, B. (1997), Ofsted criteria and the nature of social understanding: A Wittgensteinian critique. *British Journal of Educational Studies*, 45, 1, pp. 22–38.

Hood, C. Scott, C., James, O., Jones, G. and Travers, T. (1998), *Regulation Inside Government* (Oxford University Press, Oxford).

Humphrey, J. (2002), Joint reviews: Retracing the trajectory, decoding the terms. *British Journal of Social Work*, 32, 4, pp. 463–476.

Joint Review (1999), *A Report of the Joint Review of Social Services in the London Borough of Haringey. November* (Social Services Inspectorate and Audit Commission, London).

Joint Review (2001), *Delivering Results, Fifth Annual Report 2000/01* (Social Services Inspectorate and Audit Commission, London).

Miller, N. (2002), The weakest link—Priorities for improvement in personal social services performance indicators (SSRG Workshop, Leicester).

NAO (2001), *Measuring the Performance of Government Departments*, HC 301 (The Stationery Office, London).

Plank, D. (2000), Performance for the people or virtual reality? Social services and modernizing local government. *Managing Community Care*, 8, 1, pp. 13–27.

Platt, D. (2002), Why stars are underrated. *Community Care* (27 June), pp. 36–38.

Pollitt, C. and Bouckaert, G. (2000), *Public Management Reform: A Comparative Analysis* (Oxford University Press, Oxford).

Smith, P. (2002), Stars in their eyes. *Health Service Journal* (1 August), pp. 10–12.

SSI (2002a), *Performance Ratings for Social Services in England 2001–2002* (Department of Health, London).

SSI (2002b), *Performance Ratings for Social Services in England, November* (Department of Health, London).

SSI (2002c), *Performance Assessment and Annual Review 2001–02: Guidance Notes* (Department of Health, London).

SSI (2002d), *A Guide to Social Services Performance 'Star' Ratings* (Department of Health, London).