



# Rankings of Public Services – effective method of performance enhancement or counter-productive distraction?

Examples from research projects in the ESRC's Public Services Programme



**Christopher Hood**  
University of Oxford  
Director, Public Services Programme



**Ruth Dixon**  
University of Oxford

## Background

**Public Service Rankings** The UK (especially England) was an early and enthusiastic adopter of rankings for public services. Especially since Tony Blair came to power in 1997, UK public services have been subject to numerous targets, performance indicators and league tables.

**What are the aims of public service rankings?** They are claimed to increase accountability, improve performance, increase competition between service providers, and by giving the public better information, improve the choice of services.

**But do they achieve these aims?** Are the indicators robust and accurate? Are they resistant to gaming and unintended consequences? How might they be used more effectively?

Several research projects from the **ESRC-funded Public Services Programme (PSP)** have explored these questions for public services in the UK and internationally. A selection of the findings follows.

Percentages of top-rated hospital trusts and local authorities rise year on year



Figure 1. Changes in NHS star ratings and local authority CPA stars over time

## Findings

**How reliable and robust?** Jacobs and Goddard (2007) showed that both star ratings and CPA scores were highly dependent on the decision-rules and weightings chosen for the composite indicators.

**Independence of factors other than performance?** McLean et al. (2007) found that CPA scores depended (negatively) on deprivation, and it was possible for local authorities to 'buy' CPA improvements by spending above government guidelines.

**Categorization errors?** Haubrich and McLean (2006) showed that the self-assessed 'ability to improve' (a component of the CPA score) in 2002 showed no correlation with actual improvements in 2003 or 04 (Figure 2). Was self-assessment an 'easy' way for authorities to boost their CPA scores?

**Were the performance improvements real?** Propper et al. (2008) showed a reduction in hospital waiting times in England as compared with Scotland at a time when the target regime is weaker in Scotland. But O'Mahony and colleagues showed that NHS star ratings did not correlate with hospital productivity (Stevens et al. 2007).

## Cases

**Star ratings** of English National Health Service Trusts were introduced in 2000/1 and **Comprehensive Performance Assessment (CPA)** of English Local Authorities was introduced in 2002.

Both rating schemes were

- composites of 100s of performance indicators and auditor judgements
- a measure of performance against national targets
- rewarded by more autonomy from central direction
- widely reported in the local and national press
- unstable—subject to changes in methodology from year to year, and of limited lifespan (NHS star ratings ended in 2005, CPA ratings in 2008, and both were replaced by different indicators)

Substantial overall performance improvements were reported on both of these rankings since their introduction, with an increase in 'top' grades occurring year-on-year (Figure 1).

For details, see:

Haubrich and McLean 2006 *Policy Studies* 27:272-293  
Harrison and Pollitt <http://www.npcrdc.ac.uk/r5.75>  
Hood, Dixon and Beeston 2008 *IPMJ* 11:298-328  
Hood 2007 *Public Money & Management* 27:95-102  
Jacobs and Goddard 2007 *Public Money & Management* 27:103-110  
James 2007 *J. Public Administration Research and Theory* 19: 107-123  
McLean et al. 2007 *Public Money & Management* 27:111-117  
Propper et al. 2008 *CMPO Working Paper* 08/205  
Rice et al. 2008 *HEDG Working Paper* 08/28  
Stevens et al. 2006 *National Institute Economic Review* 197:80-92

No correlation between local authorities' own predictions of improvement and their actual subsequent improvement

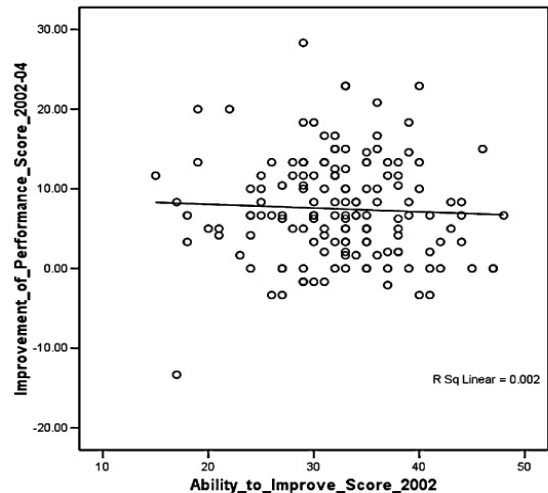


Figure 2. English local authorities' self-assessed 'Ability to Improve' in 2002 vs performance improvement between 2002 and 2004 (Haubrich and McLean 2006).

## Discussion

**Were the ratings fit-for-purpose?** Evidence that English NHS Star Ratings and local authority CPA Ratings are 'fit for purpose' is ambiguous. Performance on some measures improved, but public satisfaction with services declined (James 2007). The PSP research reported here highlights problems with these ratings.

**Where will ratings go from here?** Some other countries appear to be taking the 'English' route. For instance, Harrison and Pollitt compared the Dutch experience of introduction of Health Service indicators 20 years after England.

**Scaling back, or more of the same?** Government rhetoric in England implies a scaling back of these performance ranking regimes. But when one ranking is removed, another appears. Star ratings were replaced in 2006 by the NHS 'Annual Health Check'. The CPA rating of local authorities is being replaced in 2009 by the Comprehensive Area Assessment covering public, private and voluntary provision of public services and, like the new health check, taking survey data of public preferences into account.

**How might rankings be improved?** Questions the PSP has explored here include: Can we develop a method of kitemarking rankings for validity and reliability (Hood, Dixon and Beeston 2008)? For international comparisons, can we use 'vignettes' to anchor national preferences (Rice et al. 2008)? Can we develop a theory of when and how to use the different types of performance indicators (Hood 2007)?