

Christopher Hood, Editor

Christopher Hood
Oxford University

Gaming in Targetworld: The Targets Approach to Managing British Public Services

Christopher Hood is the Gladstone Professor of Government and a fellow of All Souls College, Oxford University. He has been the director of the Economic and Social Research Council's research program on public service quality, performance, and delivery since 2004.
E-mail: christopher.hood@all-souls.oxford.ac.uk.

To what extent did the extensive system of managing public services by targets, introduced by Tony Blair's New Labour government in the United Kingdom in 1998, reproduce the classic gaming responses associated with the Soviet Union and other centralized performance-setting systems? Combining evidence from documentary sources and interviews with high-level officials in the Whitehall bureaucracy, the author suggests that the three classic types of target gaming can be identified in this public management regime. However, the central managers of the target regime did not put substantial resources into checking performance data, took reported performance gains at face value, and had no coherent antigaming strategy.

A decade or so after the Soviet system of economic management through targets had been abandoned, a distant cousin of that approach emerged under Tony Blair's Labour government in the United Kingdom for managing the performance and delivery of public services. Building on initiatives that had been applied to subsidiary delivery agencies by previous governments, the newly elected Blair government introduced more than 300 headline performance targets applying to all government departments in 1998. The targets were linked to agreed-upon budgetary allocations with Her Majesty's Treasury, Britain's all-powerful central coordinating department, and applied to everything from local bus reliability to the staffing of the armed forces and the conduct of foreign policy. Each of the headline targets negotiated with the Treasury was accompanied by a larger set of performance indicators, and central government departments, in turn, set more detailed targets—or key performance indicators—for the delivery organizations for which they were responsible. For instance, in 2004, 10 top-level targets applying to the Health Department in England were translated into some 300 lower-level targets for the various public sector health-delivery organizations for which that department was responsible, and six top-level targets applying to the Education Department were translated into 90 “conditions” for each of the 24,000 public schools in England.

Some element of terror, too, accompanied this targets regime, particularly in England. Of course, it fell a long way short of the Soviet Union's gulags and death squads, though in the Soviet Union, state-owned enterprise managers seemed to become relatively secure in their jobs after Stalin's death, and particularly during the Brezhnev era (Braguinsky and Yavlin-sky 2000, 28, 46, 83). But some of the English public service targets, notably in health, were commonly referred to in the bureaucratic vernacular as “hanging” targets or “P45” targets (P45 is the official form you get from your employer in the United Kingdom when you are fired from your job). Starting in 2001, public sector hospitals and other public health-delivery organizations in England were given “star ratings” according to their performance on targets and other indicators, and health managers whose organizations lost their stars or never gained any could expect to be fired. And at the same time, a crack staff unit, the Prime Minister's Delivery Unit (PMDU), was set up in Whitehall's high command, drawing on practices adopted in the Education Department during the Blair government's first term. The PMDU, comprising 35 or so high-level staff members and reporting directly to Tony Blair, was set up to closely monitor reported performance on the 20-odd most politically important public service targets, and it devoted much of its energy to cajoling, persuading, and finding ways to help the laggard performers reach their targets. Tony Blair himself held “stocktakes” to assess progress on the key public service targets every two or three months, and although these meetings were often cancelled, they nevertheless represented an unprecedented level of prime ministerial attention to public service performance data. Every governmental department was subjected to an elaborate reporting cycle, and its higher echelons had to learn a new and daunting bureaucratic vocabulary of milestones, trajectories, monthly reports, and priority reviews.

Welcome to public service “targetworld,” mid-2000s U.K. style. As the country that took the centralized target approach to public service management further

than any other in recent times (and ranked higher than the United States in all six of the World Bank's governance rankings in 2004; see World Bank 2005), the United Kingdom seems to offer a compelling laboratory case for students of public management in general and of target systems in particular. Some of the reported performance improvements achieved under the system seem almost as dramatic as the Soviet Union's rapid industrialization when the Western capitalist world was stagnating in the slump of the 1930s. For instance, those improvements included a reduction in the number of patients waiting 12 months or more for surgical operations in English public hospitals, from more than 40,000 in 2001 to fewer than 10,000 in 2003 (U.K. Treasury/Cabinet Office 2004, 8); a halving of the number of asylum applications between 2002 and 2003 (U.K. Treasury/Cabinet Office 2004, 9); and a two-thirds reduction in the number of people living on the streets during the four-year period ending in 2002.

However, three crucial questions about these reported performance improvements remain. First, how much of the improvement is attributable to the targets, and how much is attributable to other changes that took place at the same time—most notably, increased spending on public services during the exceptionally flush years of the late 1990s and early 2000s? (For instance, Britain's National Health Service plan for 2000 aimed to raise spending on public health care provision by no less than 35 percent in real terms over the four years through 2004–05, and real spending per pupil for school education in England and Wales rose by roughly one-third in real terms between 1997 and 2005.) Second, to what extent do the reported performance improvements reflect real underlying improvement? And third, how far did the system manage to avoid some of the well-known dysfunctions of performance management through targets and terror? What were the strategic responses to the imposition of targets by those who were exposed to them, and what antidotes did the central managers use to limit gaming responses? Did those central managers repeat the Soviet experience or transcend it?

These questions are crucial in light of what we know about gaming and strategic behavior surrounding target systems in many other contexts, notably, socialist planning systems, accounting systems, corporate management, wartime production regimes, and economic and financial management. The debate between those who favor target systems and those who favor other approaches—such as continuous improvement—is a central issue in performance management. Scholars of target systems have identified at least three major types of gaming and strategic behavior surrounding targets. One is the much-discussed *ratchet effect*, whereby the expected tendency of target setters to fix next year's targets as an incremental advance

over last year's results causes the managers of production units to restrict performance to well below their production-possibility frontier (see Bain et al. 1987; Brown, Miller, and Thornton 1994, 93; Kornai 1992). A second type of behavior is the equally well-known *threshold effect*, whereby a uniform output target applying to all units in a system gives no incentive to excellence and may indeed encourage top performers to reduce the quality or quantity of their performance to just what the target requires (Brown, Miller, and Thornton 1994, 94). A third type of behavior consists of output distortion or the manipulation of reported results—"hitting the target and missing the point," as it was characterized in a telling phrase about health care performance coined by one senior U.K. civil servant. So how far did Tony Blair's brave new public service targetworld manage to avoid these much-discussed types of gaming?

Some Evidence

Fragmentary data and all of the familiar problems of identifying causality make it hard to answer such questions definitively, but three things can be said with some confidence. First, all of the reported performance improvements illustrated previously can by no means be attributed plausibly to the sharply increased levels of public spending on public services that occurred under the Blair government. There are at least two compelling pieces of evidence that support such a conclusion. One is that some sharp increases in reported performance took place in public service domains in which spending had not greatly increased, as in the case of the much-reduced wait times for treatment in hospital emergency rooms. The other piece of evidence comes from a comparison of reported performance for health care, where England was subject to target-and-terror regimes but other parts of the United Kingdom were not. Although funding was increasing across the United Kingdom at broadly similar rates during the early 2000s (see Alvarez-Rosete et al. 2005, 946–47), the greatest reported performance increases came in England, the only part of the United Kingdom that was subject to a target-and-terror regime, at least in the first few years of that regime (Alvarez-Rosete et al. 2005, 949; see also Bevan and Hood, forthcoming). Such evidence is admittedly at the balance-of-probability level, but it strongly suggests that targets made a marked difference in reported performance, at least in some important cases and in the early period after their introduction.

The second thing that can be said with some confidence is that the reported performance data are far from problem free. Certainly, such data did not seem to command widespread confidence among the public and the media: An official attitude survey conducted in 2005 suggested that only 37 percent of U.K. respondents thought government statistics are generally

accurate, only 14 percent thought government use official figures honestly, and the least trusted out of six statistical series were public hospital waiting list figures (ONS 2005). Reports by public audit and oversight bodies painted a less stark picture but still indicated some nontrivial problems with the performance data. For example, in the domain of publicly provided health care the following discrepancies were detected:

- Two studies found a gap of some 30 percent between levels of performance over waits in hospital emergency rooms in England as reported by providers and obtained from independent surveys of patients (U.K. Commission for Health Improvement 2004; Healthcare Commission 2005).
- In studies of an eight-minute response time target for ambulances dealing with Category A calls (life-threatening emergencies), there were large and unexplained variations in the proportion of calls logged as Category A, and ambiguity over when the clock started (Bird et al. 2005; Public Administration Select Committee 2003, 18).
- A study by the Commission for Health Improvement (2003) found evidence that in one-third of ambulance organizations, response times for Category A calls had been “corrected” to less than eight minutes in ways that could not be readily explained (see Bevan and Hood, forthcoming).
- For other hospital waiting time targets, in 2001, the National Audit Office reported evidence that nine organizations had inappropriately adjusted their waiting lists, and a 2002 Audit Commission study found reporting errors in at least one indicator for almost half of the 41 organizations it studied and deliberate misreporting at three of them (U.K. Audit Commission 2003).

Such evidence suggests that we cannot dismiss the possibility there may have been a substantial disparity between reported and actual performance in cases in which target regimes rested on data reported by the organizations that were subject to those regimes.

Third, there is evidence of nontrivial amounts of gaming of all three types identified earlier (i.e., ratchet effects, threshold effects, and output distortions). The documentary sources of such evidence come from reports by public audit bodies, evidence gathered by parliamentary committees, and other reports by professional bodies, some of which have been referred to earlier. For instance, in the health care system, the audit exercises on target gaming that had been carried out as of the time of writing indicated the existence of some degree of “creative compliance,” particularly over waiting time targets. Surveys by the British Medical Association found evidence of widespread storming

(drafting in of other medical staff and cancellation of other operations) during the period that emergency room waits were reported for target purposes in 2002 (BMA 2005; Mayor 2003). The semi-independent Commission for Health Improvement (2003) found evidence that patients were often required to wait in lines of ambulances outside emergency rooms until the hospital in question was confident that the patient could be seen within a four-hour waiting target; in some cases, hospitals responded to the target that patients had to be admitted to a hospital bed within 12 hours of emergency admission by putting “beds” into hallways and turning gurneys into beds by removing their wheels (U.K. Commission for Health Improvement 2002, para. 3.19).

Another piece of evidence can be found in the results of intensive interviews with 89 U.K. upper-level central government officials concerned with managing the target system, a study that was conducted by the author, together with Steven Kelman of Harvard University’s John F. Kennedy School of Government, in 2004 and 2005. The respondents were taken from two central agencies, the Treasury and the PMDU, and the London-based spending departments responsible for health, education, employment and welfare, and crime and security. All three types of gaming identified in the literature on Soviet and other targets—ratchet effects, threshold effects, and various forms of output distortion—were identified by these interviewees in the British public service targets regime, albeit to highly varying extents.

Ratchet effects and allied forms of gaming were observable in the way that some spending departments were said to play the targets game to their advantage by negotiating undemanding targets that would be hard to miss, linked in some cases with a reluctance to share performance information with the central agencies. In such conditions, target setting could turn into a poker game between the Treasury and those departments. For instance, in one dramatic case in which a key department had unambiguously failed to meet a performance target of central political importance to the Chancellor of the Exchequer, did the Treasury stick to the same target for the next spending round, thereby risking deeper political embarrassment if there was another failure, or quietly lower the bar, in effect rewarding failure? An example of the latter was the quiet abandonment of a target for reducing the incidence of smoking during pregnancy in 2002 after its reported incidence had gone up rather than down. Still, the Treasury had to stick to its guns in at least some cases of target failure to keep the system credible within Whitehall. Respondents from the Treasury and the PMDU identified two or three out of the 20 or so U.K. spending departments as especially adept at negotiating targets that involved little or no stretch, and some respondents pointed to targets that were

virtually unmissable, as in the case of a long-term cancer-reduction target that was almost certain to be met as a result of decisions to quit smoking that had been made as part of a social trend that had begun a decade or more earlier.

Threshold effects produced by gaming were identified by interviewees in their perceptions of the behavior of actors further down the delivery chain. Examples included (1) hospital emergency rooms that were set a blanket four-hour waiting target, with no incentives to have any patients wait for less than four hours, and (2) schools that were set pupil-attainment targets on test scores, leading teachers to concentrate on a narrow band of marginal students who were close to the target thresholds and to give proportionately less attention to those at the extreme ends of the ability range or to aspects of education beyond preparing students for those particular tests that figured in target regimes. Interviewees varied widely in the extent to which they emphasized the threshold gaming problem, but for a significant minority of them, it appeared to be highly significant. For instance, one PMDU respondent explained how she had chosen to ignore the official targets applying to the very poorly performing service for which she was responsible and instead set an unofficial “guideline,” which was deliberately not called a target in order to avoid threshold effects or politically embarrassing failure.

Respondents also identified a range of output distortion effects, particularly in the health care domain but also in other cases, such as treatments or training commenced but not completed in situations in which the targets only measured commencement (as in the case of a drug treatment target that sought to double the numbers in treatment). In health care, at least 20 percent of general practitioner organizations in England were at one point perceived to be meeting a target that patients should be able to see a doctor within 48 hours (or a primary care professional within 24 hours) by not allowing anyone to book an appointment more than 48 hours in advance. That outcome was said to be far from the intention of those who had framed the target, caused political embarrassment to the prime minister when it was aired during the 2005 general election campaign, and was admitted to be a problem after that election by a junior health minister (Waugh 2005).

Interviewees varied widely in the extent and type of gaming they identified. In most cases, they saw the extent of target gaming as either unknown to them or of minor significance, and this response came particularly from younger respondents who had little experience with frontline management. In some cases, respondents echoed George Washington’s Plunkitt’s famous distinction between “honest” and “dishonest” graft (Plunkitt 1905) by separating “cheating” from

“gaming” or acceptable from unacceptable gaming. One respondent compared the boundary between the two to the distinction between tax avoidance and tax evasion, with outright falsification or making up of numbers counted as cheating but creative classification or interpretation (or creative compliance) considered gaming.

Table 1 makes a simple distinction between four types of gaming or strategic behavior in relation to target systems by distinguishing cases in which output remains essentially unchanged from cases in which output alters and by distinguishing cases in which performance data are “spun,” or creatively interpreted, from those in which data are invented or made to disappear. Of the four types of target gaming, type 4 (fictional performance data are reported) was generally seen by interviewees as cheating. Type 2 (performance improvements are reported without underlying changes in output by storming or other methods) was generally seen as an unacceptable form of gaming. Type 1 (reportable performance improvements are contrived by redistributing service from one set of users to another—for instance, by relocating ambulances from rural to urban locations so that urban response times improve at the expense of longer rural response times) constituted a class of responses that interviewees saw as acceptable in some cases but not in others. Type 3, by contrast (performance problems are made to disappear by quietly dropping the targets—a tactic that tended to apply much more to targets concerning central government than local government), was seen by interviewees as a fact of life in bureaucratic politics. One respondent with more than 30 years of experience in the employment bureaucracy thought the balance had shifted during that time from cheating, in the sense of simply making up the numbers that the central office was assumed to want (that is, type 4 in table 1), which the interviewee said was rampant during the 1970s, to gaming, in the sense of creative categorization to maximize points in the target system (somewhere on the boundary between types 1 and 2). One of the old hands in the health bureaucracy made a similar point about the shift from reporting fictionalized numbers to creative interpretation over a decade, implying that gaming had moved northwest in the array of types presented in table 1.

Although respondents varied in the estimates they gave about how much gaming was going on, a minority estimated the rate of gaming as substantial. As noted previously, some respondents saw 20 percent of primary care health trusts as gaming the 24/48-hour access target, and most interviewees in the health care domain saw the gaming of emergency room waiting time targets as more than an occasional phenomenon. One respondent thought that a majority of local offices were gaming a key welfare target by creatively classifying the clients they served to maximize their

Table 1 Four Types of Target Gaming

		Underlying Performance in Provision of Service	
		Alters	Remains unchanged
Performance Data	Creatively interpreted, contrived, or spun	(1)	(2)
		Example: Change of provider focus to meet targets Result: Data showing the target as being met, but with redistribution of service among users	Example: Storming to meet targets at one point in time Result: Data showing the target being met, but with no improvement in real performance over time
	Invented, dropped, or not provided	(3)	(4)
		Example: Quiet removal of target after a performance drop Result: Data showing a failure to achieve target disappears	Example: Reported results are more or less fabricated Result: Data showing the target as being met, but with no improvement in real performance

point scores, as some clients brought more points in the target system than others. Such responses are hard to assess, but they indicate that a small number of interviewees thought that the gaming problem was far from insignificant.

Brave New World, or Back to the Future?

If target gaming of each of the three classic types is observable in this system, did the United Kingdom's public service targetworld have antidotes to the target-gaming problem that were not available to the Soviet Union or other earlier target systems?

To some extent, this may have been the case. It is true that some of the standard responses took the form of ever more elaborate books of rules and tightening of data definitions to close down the creative interpretation of targets (a marked feature of the health care targets), modification of targets or key performance indicators at the next level down from targets to make them harder to game (a marked feature of the job-entry target regime), and sporadic investigations by public auditors or other oversight bodies, some of which have been quoted earlier. Some interviewees stressed the efficacy of threats of draconian punishment for managers caught abusing the system—the classic strategy of killing one admiral to encourage the others, in Voltaire's (1759, 224) famous satirical comment on 18th-century British naval administration—though such views are hard to triangulate with other data because there is no systematic evidence of an increase in firings or other forms of discipline among public service managers.

Many of the 89 interviewees professed to seeing audit and oversight as a key antidote to target gaming, and some stressed the efficacy of “mystery shoppers” as a way of checking service performance, notably in health care and job entry. But such optimistic views about the effectiveness of mystery shoppers as an antidote to gaming were not shared by all respondents; several stressed the high predictability of the

process in practice. Moreover, there was no perceptible increase in resources devoted to audit bodies to tackle gaming by looking carefully behind the reported numbers, and in health care, the move was, if anything, in the opposite direction (see Bevan and Hood 2004).

Indeed, a possible type of gaming that is not fully captured in table 1 might have consisted of an eagerness by the central managers of the system to accept “good news” performance data at face value without putting substantial resources into probing those numbers. After all, given all that is known about gaming problems in target systems, we might have expected those managers to design a system to tackle the gaming problems from the outset. As it was, the responses to gaming problems tended to come well down the line, and even then it is debatable how effective they were. In health care, for instance, waiting time performance in key targets was assessed using the data the health care organizations reported rather than the results of patient surveys, in spite of the large discrepancy between those two sets of numbers. Patient survey data fed into a balanced scorecard that was of lesser importance than the key targets, yet if patient survey data had been used instead of organizationally reported waiting times to calculate hospital star ratings, at least one politically sensitive hospital trust would not have achieved a top rating. Instead of putting more resources into verifying numbers on the ground, the audit system stressed remote statistical analysis of the numbers reported. A unit for verifying health care performance data was finally set up in 2005 and much trumpeted as an “independent” body, though, in fact, it was set up as a trust reporting to the Department of Health, which controlled appointments and procedures—subtly different from the role of an “office of performance data,” as conceived in the United States by Robert Behn (2001, 204).

Why was there no real attempt to check such data properly from the start? The slow and half-hearted

approach to developing independent verification of performance data might itself be interpreted as a form of gaming by the central managers (like the famous English admiral, Horatio Nelson, who put a telescope to his blind eye to avoid seeing a signal he did not want to obey). Or, in a more innocent interpretation, it might simply represent the priorities of a central government in a hurry to show results without imposing extra regulatory burdens on public services.

We do not have the evidence to assess which of these two interpretations is correct, and indeed, hard evidence on that issue may never be available. But whichever explanation is correct, given all that is known about gaming in target systems of the Soviet type, it is remarkable that so little was done from the outset, not only to verify and check reported performance data in some crucial areas but also to avoid some of the very well-known pitfalls of target systems (such as 100 percent achievement of targets, as originally set for emergency room wait times, and the setting of common targets for different organizations, as in the school system). It seems like a case of those who refuse to learn lessons from historical experience being forced to repeat that experience.

If all of the antidotes to target gaming discussed so far—tightening the rules and data definitions, refining the targets, conducting audit investigations, threatening to discipline errant admirals—would be familiar enough to Gosplan veterans,¹ three approaches were arguably rather less so. One was an increasing reliance in some areas on third-party collection of data in the form of user satisfaction surveys rather than data directly collected by the organizations subject to the targets. A second (admittedly a technologically advanced version of a far older technique) was the development of information system architecture that limited creative responses to data entry—for instance, by barring multiple entries or preventing a job placement from being recorded unless there was a corresponding entry in which the client came off welfare benefits. And a third approach was a real attempt in some policy domains, particularly in the crucial case of health care at the end of the Blair government's second term, to escape the ratchet effect by essentially declaring victory at a particular point of performance and thereafter turning targets linked to the allocation of funds from the Treasury into permanent threshold standards policed by quality overseers (see U.K. Treasury/Cabinet Office 2004, 19; for unsuccessful Soviet attempts to prevent ratchet effects in the target system during the 1970s and 1980s, see Hewett 1988; Nove 1986).

What seemed to be lacking, though, was any coherent antigaming strategy. Several of the more reflective and managerially minded interviewees saw target gaming as the product of “blaming, bullying, non-

transparent [organizational] cultures” in public-service-delivery organizations, but many of the antidotes identified by those respondents tended to focus on limiting opportunities to game the system (by low-trust measures to check up on or constrain the gaming options) rather than changing motivation by active measures to reduce bullying and blaming. Indeed, many saw the drift of organizational change as going in the opposite direction.

Conclusion

Upon returning from his visit to the Soviet Union in 1919, Lincoln Steffens famously observed, “I have seen the future and it works” (1938, 463). Should we draw the same conclusion today from the United Kingdom's target system for managing public service performance, which reached its peak during the Blair government's second term? Only up to a point, it seems. As noted earlier, the United Kingdom performed strongly in the 2004 round of the World Bank's general governance rankings, but on many of the public-service-specific international rankings—for educational attainment, health, and crime, for instance—it did not score uniformly in the top third of the leading 12 or so countries of the Organisation for Economic Co-operation and Development. Moreover, both the audit data mentioned and the responses of the 90-odd Whitehall insiders interviewed for this study indicate that the three classic gaming problems associated with target systems all appeared to some extent, even though reported improvements were certainly dramatic in some cases. But if Tony Blair's United Kingdom was indeed unconsciously repeating Soviet history as a public-service targetworld, what remains to be seen is whether that target system will last for anything like the 60 years that the Soviet system survived. At the time of this writing, the target approach, the centerpiece of Blair's second term, was quite widely perceived as “last week's salad,” and the hot debate at the heart of New Labour in Britain was how far choice—linked to performance indicators to provide comparative information for users—could replace Soviet-type targets for public services. But to the extent that choice cannot fix everything—particularly in public services with sharp capacity limits or vulnerable consumers (such as Alzheimer's patients looking for care homes)—the question of how to fine-tune or game-proof target-and-performance indicator systems remains important.

Note

1. In the Soviet Union, the State Planning Committee (*Gosudarstvennyi planovyi komitet*, commonly known by its acronym, *Gosplan*) was primarily responsible for devising and monitoring five-year plans and annual plans. The name was changed in 1948, but the acronym was retained.

References

- Alvarez-Rosete, A., G. Bevan, N. Mays, and J. Dixon. 2005. Effect of Diverging Policy across the NHS. *British Medical Journal* 335: 946–50.
- Bain, J., M. Keren, J. Miller, and J. Thornton. 1987. The Ratchet, Tautness and Managerial Behaviour in Soviet-Type Economies. *European Economic Review* 31: 1173–1201.
- Behn, R. 2001. *Rethinking Democratic Accountability*. Washington, DC: Brookings Institution Press.
- Bevan, R. G., and C. Hood. 2004. Targets, Inspections and Transparency. *British Medical Journal* 328: 598.
- . Forthcoming. What's Measured Is What Matters: Targets and Gaming in the English Public Health Care System. *Public Administration*.
- Bird, S. M., D. Cox, V. T. Farewell, H. Goldstein, T. Holt, and P. C. Smith. 2005. Performance Indicators: Good, Bad, and Ugly. *Journal of the Royal Statistical Society* 168(1): 1–27.
- Braguinsky, S. V., and G. Yavlinsky. 2000. *Incentives and Institutions: The Transition to a Market Economy in Russia*. Princeton, NJ: Princeton University Press.
- British Medical Association (BMA). 2005. *BMA Survey of A&E Waiting Times*. London: British Medical Association.
- Brown, P., J. Miller, and J. Thornton. 1994. The Ratchet Effect and the Coordination of Production in the Absence of Rent Extraction. *Economica* 61(241): 93–114.
- Hewett, E. A. 1988. *Reforming the Soviet Economy: Equality versus Efficiency*. Washington, DC: Brookings Institution Press.
- Kornai, J. 1992. *The Socialist System: The Political Economy of Communism*. Oxford, UK: Clarendon Press.
- Mayor, S. 2003. Hospitals Take Short Term Measures to Meet Targets. *British Medical Journal* 326: 1054.
- Nove, A. 1986. *Socialism, Economics and Development*. London: Allen and Unwin.
- Plunkitt, G. W. 1905. *Plunkitt of Tammany Hall: A Series of Very Plain Talks on Practical Politics*. Recorded by W.L. Riordan. New York: McLure, Phillips.
- Steffens, L. 1938. *The Letters of Lincoln Steffens*. Vol. 1. New York: Harcourt Brace.
- United Kingdom. Audit Commission. 2003. *Waiting List Accuracy*. London: Her Majesty's Stationery Office. www.audit-commission.gov.uk [accessed April 3, 2006].
- United Kingdom. Commission for Health Improvement. 2002. Report on the Clinical Governance Review on Surrey and Sussex Healthcare NHS Trust. London: Her Majesty's Stationery Office. www.healthcarecommission.org.uk [accessed April 3, 2006].
- . 2003. *What CHI Has Found in Ambulance Services*. London: Her Majesty's Stationery Office. www.healthcarecommission.org.uk [accessed April 3, 2006].
- . 2004. *What CHI Has Found in Acute Services*. London: Her Majesty's Stationery Office. www.healthcarecommission.org.uk [accessed April 4, 2006].
- United Kingdom. Healthcare Commission. 2005. Patient Survey Programme 2004/2005. Emergency Departments: Key Findings. www.healthcarecommission.org.uk [accessed April 4, 2006].
- United Kingdom. Her Majesty's Treasury/Cabinet Office. 2004. *Devolution Decision-Making: 1—Delivering Better Public Services: Refining Targets and Performance Management*. London: Her Majesty's Treasury.
- United Kingdom. National Audit Office. 2001. *Inappropriate Adjustments to NHS Waiting Lists*. London: Her Majesty's Stationery Office. HC452. www.nao.org.uk/publications/nao_reports/01-02/0102452.pdf [accessed April 3, 2006].
- United Kingdom. Office of National Statistics (ONS). 2005. No Change in Public Confidence in Official Statistics. News release, September 22. www.statistics.gov.uk/pdfdir/pco0905.pdf [accessed April 3, 2006].
- United Kingdom. Public Administration Select Committee. 2003. *On Target: Government by Measurement*. London: Her Majesty's Stationery Office. HC62-I.
- Voltaire, F.-M.A. 1759. *Candide*. Vol. 48 of *Les oeuvres completes de Voltaire*. Oxford, UK: Tylorian Institution, 1980.
- Waugh, P. 2005. Minister Admits Failure over GP Bookings. *Evening Standard*, September 7, 14.
- World Bank. 2005. Governance Research Indicator Country Snapshot (GRICS): 1996–2004. <http://info.worldbank.org/governance/kkz2004/index.htm> [accessed April 3, 2006].

Copyright of *Public Administration Review* is the property of Blackwell Publishing Limited and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.