

Judicial accountability in the US state courts Measuring court performance*

Richard Y. Schauffler**

Currently state courts in the United States are reexamining the issues of performance measurement and accountability. In part a response to competition for scarce budget resources, in part a response to the growing politicization of judicial decision-making and judicial selection, this reawakening of interest in and implementation of performance measurement is also a reflection of struggling with the increasing complexity of modern court management and the need for some objective measurement to inform court managers, the public and partners in government about how well the courts are managed. Here we examine the current efforts at performance measurement in the state courts, situated in a global and historical context.

What explains the resurgence of interest in court performance measurement as a particular form of court management reform? Why are the US state courts of interest to Europe, and EU Member States in particular? The experience of US state courts might be a useful reference point for several reasons. The current impetus for court performance measurement is particular to the US, but it is also an instance of public management reform that can be understood in a comparative framework. What is taking place in the US is not of interest because developments in the US will necessarily, through some unspecified process of 'globalization', directly influence developments in Europe. As Pollitt and Bouckaert point out in their model of public management reform, global forces operate in complex ways through intervening variables.¹ The value of understanding the developments arising out of the federation of the fifty states of the US is that these developments may prove illuminating for understanding developments unfolding with the expansion and strengthening of the EU as a supranational polity. Just as the EU Member States enjoy formal institutional autonomy among their sovereign national legal and judicial systems, so do the fifty US states. Just as the EU Member States find themselves at varying levels of institutional maturity and capacity, so do the fifty US states. While obviously the fifty states share a common legal framework, the challenges of institutionalizing performance measurement in this decentralized and diverse US environment is perhaps more relevant to the EU than it might appear at first glance.

* I am indebted to my colleagues at the National Center for State Courts for our extensive collaborative work on court performance measurement, notably Brian Ostrom, Ingo Keilitz, Dan Hall, Tom Clarke, Fred Cheesman, and Charles Ostrom. While the inspiration for this article comes from our theoretical and practical work together, the views expressed here are the sole responsibility of the author.

** Richard Y. Schauffler is Director of Research Services at the National Center for State Courts, 300 Newport Avenue, Williamsburg, VA 23185, USA (rschauffler@ncsc.dni.us). At the NCSC, he directs the Court Statistics Project, which collects and analyzes state court data to develop a national perspective on the work of the state courts. He also works on the CourTools to assist states and courts with implementation of performance measures. His third area of concentration is workload assessment projects for judges, clerks, staff and public defenders. Mr. Schauffler holds a bachelor's degree from the School of Criminology, University of California at Berkeley, and an M.A. in sociology from the Johns Hopkins University.

1 C. Pollitt & G. Bouckaert, *Public Management Reform: A Comparative Analysis*, 2000, pp. 24-38.

1. The global context

Over the last two decades, courts around the world have been struggling with performance measurement as part of broader public sector and judicial reforms. For developing countries and post-Soviet countries of Central and Eastern Europe, these attempts have been made in the context of democratization and the penetration of Western investment, which insists upon a more extensive and reliable rule of law to protect its investments. In Western Europe, the focus on quality derived from reasons similar to those in the US: renewed emphasis – brought on by global economic crisis beginning in the mid-1970s, the privatization and contracting out of public services of the 1980s, and the neoconservative political pressures of the 1980s and 1990s² – on defining the appropriate type and quality of public services. In the United Kingdom, this movement developed under the name New Public Management, which was embraced to varying degrees (and with varying but mostly inconclusive results) in many countries around the world.³ In the US, given the fragmented nature of state politics, performance measurement for courts has not been part of a concerted joint effort across states, but rather a project undertaken or not according to the priorities of state and local level judicial leadership.

The comparative framework provided by Pollitt and Bouckaert (2000) for analyzing public management reform at the national level as instigated by executive government can help illuminate the same issues for the judiciary.

That heuristic model contemplates the following variables:

A. Socio-economic forces

- B. Global economic forces
- C. Socio-demographic change
- D. National socio-economic policies

E. Political system

- F. New management ideas
- G. Party political ideas
- H. Pressure from citizens
- I. Elite perceptions of what management reforms are desirable
- J. Elite perceptions of what management reforms are feasible
- K. Chance events, *e.g.*, scandals, disasters

L. Administrative system

- M. Content of reform package
- N. Implementation process
- O. Reforms actually achieved⁴

The relationships among these variables are diagrammed in the original, but for the purposes of the present discussion it is sufficient to enumerate them and discuss those relevant for under-

2 See C. Pollitt, *Managerialism and the Public Services: The Anglo-American Experience*, 1990, pp. 28-49 for a more detailed description of this trajectory.

3 For a brief description of the basic elements of NPM, see R. Gregory, 'Accountability in Modern Government', in B. G. Peters & J. Pierre (eds.), *Handbook of Public Administration*, 2003, pp. 563-567, and C. Hood & M. W. Jackson, 'The New Public Management: A Recipe for Disaster', 1991 *Canberra Bulletin of Public Administration* 64, pp. 16-24. For a more detailed and historical discussion of the economics and politics of NPM, see L. E. Lynn, Jr., *Public Management: Old and New*, 2006, pp. 104-156. NPM is discussed in the context of a ten-country comparison in C. Pollitt and G. Bouckaert, *Public Management Reform: A Comparative Analysis*, 2000.

4 From C. Pollitt & G. Bouckaert, *Public Management Reform: A Comparative Analysis*, 2000, p. 26 Figure 2.1.

standing the current court management reforms underway in the US state courts. Prior to that, it is necessary to provide some historical and descriptive background on the evolution of the state courts in the US.

2. The national context

In order to understand developments in court performance measurement in the state courts, it is necessary to situate it in the context of the political and economic structures of state government in the US. At the outset, it is important to distinguish between the federal court system and the state court system. The federal courts are part of the national government, and are funded, organized, and administered at the national level. Tensions exist within the system between the jurisdictions between these two court systems (some matters are exclusive to one or the other, some are overlapping), the role of federal and state courts with respect to finality of state court decisions, federal preemption of state law, and the control over administration of the state courts and the bar. These issues, while legally important, are not the heart of the matter here; it is sufficient for our purposes to clearly note that federal courts are not addressed here.

In some sense, the term ‘state courts’ may itself be somewhat misleading, since in some states it refers to a coherent set of courts functioning as a judicial branch in concert, while in other states it refers to local-level courts that are geographically located within the same state boundaries, but which have little or nothing to do with each other. The reasons for this are explained below.

2.1. The state judicial branch

The state courts in the United States are situated within a government that is constitutionally structured with three branches: the executive, the legislative and the judicial. In theory, the first two branches govern, while the third adjudicates. If each branch adheres to its role, this structure provides a means through which each branch regulates the behaviour of the other branches and constrains their ability to act outside their constitutionally mandated role, thus preventing the abuse of power. In the popular vernacular, this system is described as ‘the system of checks and balances’.

When applied to the state and local government, however, this simple but elegant model does not fit reality as well as it might at the national level. The state courts evolved as local institutions at the city or county level, funded by the local government and originally adjudicated by lay judges on the basis of local custom and law. As such, the trial courts (courts of first instance) were and to some extent remain embedded in local politics and highly invested in local practices. The courts of last resort (usually referred to as state supreme courts) and intermediate appellate courts functioned as statewide institutions reviewing the work of the lower courts upon demand. The ongoing attempt throughout the nineteenth and twentieth centuries to create a state-level judicial branch of government from the numerous county or city-level courts is at best an ‘unfinished reform’.⁵

2.2. State court budgets

About one-third of the fifty states now fund their courts through a central state budget; the remaining states fund their courts primarily through local funding at the county or city level.⁶ In

⁵ R. Tobin, *Creating the Judicial Branch: The Unfinished Reform*, 1999, p. 3.

⁶ D. Rottman et al., *State Court Organization 2004*, 2004, pp. 83-96.

many states this lack of budget control creates very weak state-level administrative offices of the courts, which exercise leadership through cooperation and persuasion, not through the leverage of resource allocation.

Regardless of the source of funds, control of the judicial branch budget further problematizes the notion of judicial branch independence. In 29 of the 50 states, the judicial branch budget is submitted to the executive branch, and is subject to negotiation with the governor prior to being submitted to the legislature.⁷ The history of control over judicial budgets is a long and torturous one; more than one state supreme court has seen its budget cut as the result of a court decision that was unpopular with its governor or legislature. At the local level, court budgets are subject to line item review and veto by the local executive or legislative branch, which constrains the ability of the judiciary to chart its own course as an institution. The issue was fully explored in Baar's book, whose title expresses the problem: *Separate but Subservient: Court Budgeting in The American States*.⁸ Although the judicial branch budgets in the fifty states range from 0.5 to 5% of the total state budget,⁹ in larger states this translates into hundreds of millions and even billions of dollars. With this growth in judicial branch budgets has come a demand by the executive and legislative branches for greater accountability.

2.3. State court politics

In most states, the judicial branch has traditionally held itself aloof from partisan politics, seeking to remain in perception and in actuality disengaged from the political fray. This distancing was seen as necessary for remaining impartial and effective arbiters of legal disputes. In recent years, however, the theory and practice of political neutrality and temperance has come under pressure from political interests (party political ideas (G) in the model). One form of this pressure comes from the legislative branch; through the practice of 'court stripping', whereby legislatures seek to remove the authority of the courts to review the legality of particular laws. For example, in the *Schiavo* case,¹⁰ Congress sought to eliminate the authority of the Florida state courts by imposing federal jurisdiction. State-level court stripping efforts occurred recently in Florida and Arizona where the legislatures sought to transfer the power to write rules of court from the state supreme courts in those states to the legislatures.¹¹

Another form of political pressure comes from political interest groups seeking to influence judicial elections or pass legislation through popular ballot initiatives to restrict the authority of courts. The developments here are quite negative, as a few highlights from 2006 illustrate:

1. Television advertisements ran in ten of 11 states with contested Supreme Court elections, compared to four of 18 states in 2000...
2. Average spending on television airtime per state surpassed \$1.6 million...
3. Pro-business groups were responsible for more than 90% of all spending on special interest television advertisements...
4. Candidates themselves went on the attack, sponsoring 60% of all negative ads...¹²

7 *Ibid.*, pp. 80-82.

8 Carl Baar, *Separate but Subservient: Court Budgeting in The American States*, 1975.

9 D. Rottman *et al.*, *State Court Organization 2004*, 2004, pp. 80-82.

10 The legal developments surrounding the right to die of Terri Schiavo are actually a series of state and Federal court cases. Case references and legal documents can be found at <http://www.euthanasiaprocon.org/legaldocumentsschiavo.html>.

11 David Rottman, 'The State Courts in 2005: A Year of Living Dangerously?', in Council of State Governments, *The Book of the States 2005*, 2006.

12 J. Sample & L. Jones, *The New Politics of Judicial Elections 2006*, <http://www.justiceatstake.org/contentViewer.asp?breadcrumb=3,570,979> (accessed 21 May 2007).

Due to the peculiar form of electoral politics practiced in the US (where political parties are mere advertising labels for individual politicians who are not bound by a party programme, and where special interests, rather than political parties, reign), the analysis of these developments problematizes both party political ideas (G) and pressure from citizens (H) in the model. Business interests frequently masquerade as citizen interests through fictive organizations and disguised spending. While the legislative and executive branches feel it is appropriate (if not required) to be responsive to this kind of political pressure, judicial branch leaders have been appropriately reluctant to allow this kind of political activity to influence the judicial branch. In this context, performance measurement is appealing as it has the potential to shift the discourse in elections from specific decisions in individual cases to more appropriate criteria.

2.4. State court structure

State courts in the United States evolved at the city or county level within the context of the system of three branches of government outlined above. As a result of the historical evolution of the country as it expanded westward from the seventeenth through the twentieth century, the structure of the courts in each state is far from homogeneous. In general, although the names of these courts vary widely along with the definitions of their jurisdiction, the courts in each state comprise the following jurisdictional levels: a court of last resort (usually but not always called the supreme court); an intermediate appellate court; and one or more levels or types of trial courts (courts of first instance). The trial court levels typically include a general jurisdiction court (typically called a superior court, district court, or circuit court) and one or more types of limited jurisdiction court (*e.g.*, town and village court, municipal court, traffic court).¹³

The distinction between the two levels of trial court can be characterized in terms of the seriousness of the cases that come before them, expressed as either the economic value of the case (for a civil case) or the severity of the crime and its sanction (for a criminal case). In some states, the courts are combined through a system of geographical regions, through which several courts form a single district or circuit served by a common pool of judges. It is safe to say that almost every variation of the above exists today, from the simplified unified structure of states like Minnesota and California (with a single court of last resort (supreme court), a single tier of intermediate appellate courts (court of appeals) and a single tier of trial courts (district court or superior court, respectively) to the still complex structures of states like New York (with eight distinct limited jurisdiction courts (courts of claims, surrogate's court, family court, district court, city court, town and village justice court, civil court of the city of New York, criminal court of the city of New York), two different general jurisdiction courts (county court and the confusingly named supreme court), two different intermediate appellate courts (appellate divisions of the supreme court and appellate terms of the supreme court) and one court of last resort (the confusingly named court of appeals). The unnecessary complexity of many of these state court structures was decried in Roscoe Pounds's famous 1906 address to the American Bar Association and cited as one of the 'causes of popular dissatisfaction with the courts'.¹⁴

As states sought to create a state-level judicial branch from these local-level institutions, they typically assigned responsibility for judicial branch governance either to the state supreme court, led by a chief justice, or to a judicial council, a small, representative body chaired by the chief justice. Either way, the state-level institution seeks to govern and administer the local-level

13 The most complete guide to the organization of the state courts in the US is *State Court Organization, 2004* a joint project of the Bureau of Justice Statistics and the National Center for State Courts, available at <http://www.ojp.usdoj.gov/bjs/abstract/sco04.htm>

14 R. Pound, 'The Causes of Popular Dissatisfaction with the Administration of Justice', 1964 *Crime and Delinquency* 10, pp. 355-371.

courts through the creating of mandatory statewide rules of court, standards of judicial administration and disciplining of judicial officers. The judicial branch is led by two figures: the chief justice of the state supreme court (appointed by the governor in thirteen states, elected in non-partisan election in four states, elected in partisan elections in two states and chosen by other means, typically the supreme court itself, in the remainder)¹⁵ and the state court administrator, who serves at the pleasure of the supreme court. However, it must be noted that the notion of a professional manager as part of the state-level administration of the judicial branch is a rather modern one: the first such position was not created until 1947 in New Jersey.¹⁶

The state court administrator serves as the chief executive officer of an administrative office of the courts, the statewide staff office for the judicial branch. The authority and span of control of that office varies widely. In those states where the judicial council/supreme court gained control over the funding of the courts through the state budget, this provides additional leverage through the distribution of these funds to local courts via the administrative office of the courts. In these states, fiscal control can also be used to create more standardized, statewide approaches to judicial administration.

This brief discussion of state court structure would not be complete without making mention of the two most influential actors in the state courts at the local level: judicial officers (*i.e.*, justices, judges, commissioners, referees, magistrates, justices of the peace, by whatever names they are known) and court executive officers (*i.e.*, court administrators or clerks of court, by whatever names they are known). Judicial officers at these various levels of state court arrive at their position in a variety of ways. For general jurisdiction trial courts, judges are elected in partisan or non-partisan elections in 27 states; appointed by the governor in 19 states; and appointed by the legislature or other means in the remaining states.¹⁷ The governor and the legislature both play a role in the authorization for and funding of new judgeships; in about 35 states, requests for new judgeships are made on the basis of a quantitative workload assessment methodology administered by the judiciary.

Within each trial court one judge is typically elected by fellow judges locally or appointed by the chief justice of the state supreme court to serve as the presiding judge (also known as the chief judge or administrative judge). This judge is ‘first among equals’ and is charged with managing the legal and judicial side of the court and co-managing the administrative matters with the court administrator. All the dynamics of competing allegiances described below are at play with respect to this position as well.

Court administrators are either elected in local partisan or non-partisan elections in 30 states, or appointed by either state-level (administrative office of the courts) or local-level (the trial court) judiciary.¹⁸ While the office of an elected clerk of court is an old tradition in many states, the notion of a professional manager hired for their management expertise is a very modern one: the first such administrator was hired by the Superior Court in Los Angeles, California, in 1957.¹⁹

The appointment/election/hiring process of these key actors itself thus poses an immediate question: To whom are judges and court administrators accountable: the governor, the chief justice of the state supreme court, the state legislature, the county, the litigants, or the taxpaying and voting public? Are they agents of a state entity – the judicial branch of state government –

15 D. Rottman *et al.*, *State Court Organization 2004*, 2004, pp. 25-28.

16 A. Aikman, *The Art and Practice of Court Administration*, 2007, p. 2.

17 D. Rottman *et al.*, *State Court Organization 2004*, 2004, pp. 33-38.

18 *Ibid.*, pp. 169-174.

19 L. Sipes, *Committed to Justice: The Rise of Judicial Administration in California*, 2002, p. 69.

or local officials? Their interest in and implementation of performance measurement is of course affected by their views on this issue. These are not merely rhetorical questions, since state courts are responding to inputs from both the state and local levels. In the discussion of the Pollitt and Bouckaert model below, we will confine our discussion to the state level.

3. Performance measurement in the state courts of the US

In the wake of the neoconservative efforts in the 1980s to privatize, downsize, and reorganize the public sector, the idea of reconceptualizing how governmental functions are organized and evaluated held some appeal across the political spectrum. An effort was made to transcend the significant but ultimately sterile debate about more or less government and to shift the discourse to answering the question: Government for what purpose? Perhaps the most popular book on the subject in the US, *Reinventing Government*,²⁰ argued for performance-based public institutions. The use of performance measurement called for two things: clarification of the purposes of each institution and definition of the appropriate measures to gauge progress toward those specific organizational objectives.

While much of this discourse in the US was about the federal government,²¹ the ideas permeated state and local government as well. States like Colorado and Connecticut undertook benchmarking efforts, as did some major cities. Indeed, by the mid-1990s, it was claimed that ‘performance measurement, the regular collection and reporting of information about the efficiency, quality and effectiveness of government programmes is arguably the hottest topic in government today’.²²

In the state courts, the impetus toward performance measurement was further stimulated by several other factors:

1. the enormous increase in cases prosecuted as part of the national ‘war on drugs’ that were overwhelming the courts;
2. renewed attention to court delay and costs of litigation;
3. the economic recession of the early 1990s, which put a serious budget squeeze on state and local budgets and illustrated how ineffective courts were at justifying their use of public dollars with objective data,²³ and
4. the low level of public trust and confidence in the courts, as reported in national and state surveys.²⁴

To reframe these elements within the framework provided by Pollitt and Bouckaert, the factors pushing the reform effort included: 1) factors from the political system (new management ideas (F)); 2) socio-economic forces (socio-demographic changes (C) in terms of increased drug use and drug-related crime), which in turn result in elite perceptions of how to respond to 3) a second set of factors from the political system (party political ideas (G) and pressure from citizens (H)

20 D. Osborne *et al.*, *Reinventing Government*, 1992.

21 National Partnership for Reinventing Government (NPR), originally the National Performance Review, was an interagency task force created in March 1993 during the Clinton presidency to reform the way the U.S. federal government works.

22 R. Nyhan *et al.*, ‘Comparative Performance Measurement’, 1999 *Public Productivity and Management Review*, no. 3, pp.348-364.

23 National Association of State Budget Officers, ‘Fact Sheet: Quick Rebound? State Fiscal Recovery Could Be Gradual, Lag National Economy 12-18 Months’, 2002.

24 See for example American Bar Association, *Perceptions of the US Justice System*, 1999; National Center for State Courts, *How the Public Views the State Courts: A 1999 National Survey*; Yankelovich, Skelly, and White, Inc., *The Public Image of the Courts: Highlights of a National Survey of the General Public, Judges, Lawyers, and Community Leaders*, 1978.

regarding drug use and drug-related crime, leading to the ‘war on drugs’); 3) a third set of factors from the political system (pressure from citizens/interest groups (H) on court delay and legal costs); 4) socio-economic forces (national and state economic conditions and policy responses (D)); and 5) additional factors from the political system (pressure from citizens (H) in the form of negative public opinion about the courts) which in turn give rise to new elite perceptions of what management reforms are desirable (I).

In this context, the National Center for State Courts initiated the development of what came to be known as the Trial Court Performance Standards (TCPS). Over three years, the Commission on Trial Court Performance Standards engaged the court community in the development of 22 standards requiring 68 measures across 5 broadly defined areas: access to justice; expedition and timeliness; equality, fairness and integrity; independence and accountability; and public trust and confidence.²⁵ Conceptually, the TCPS were aimed at evaluating the performance of the court as an organization, not the performance of individual judicial officers per se. The point of reference was those who use the court, and the focus was on how to improve services to the public. The TCPS were first published in 1990 and endorsed by all the key national court organizations (Conference of Chief Justices, Conference of State Court Administrators, National Association for Court Management and American Judges Association).²⁶ Again, referring to the model, these management reforms were endorsed by judicial leadership as desirable (I).

The purpose of securing endorsements from national organizations is perhaps obvious: faced with a federation of state court systems, any effort to create a coherent national project must gain acceptance from the leaders of those state and local courts. While the members of those associations are not bound to carry out the resolutions or policy guidelines developed at national meetings, there is nonetheless important symbolic value and legitimacy conveyed by being able to refer to these actions.

As a sign, however, of how intimidated the court community was by the notion of performance measurement, the TCPS were published with an explicit disclaimer that the measures were only to be used for a court’s internal management. This message was not lost on the states. The Judicial Council of California, for example, adopted the TCPS as one of its Standards of Judicial Administration with an explicit statement that the standards are ‘for the purposes of internal evaluation, self-assessment, and self-improvement. They are not intended as a basis for cross-court comparisons’.²⁷ Whether this was the price that had to be paid for embracing the TCPS in that state or any other can be debated, but the point here is simply that this was indicative of the fear of data-driven performance evaluation that was widespread among the state courts.

For a variety of reasons, shortly after it was endorsed the movement toward performance measurement in the state courts lost momentum at this time. As summarized by one court management expert, ‘At an intellectual level, the Trial Court Performance Standards changed the debate, and changed the perception about the value of data. Regrettably, they did not change operations in more than a few courts’.²⁸ Several factors contributed to the inability of state courts to institutionalize performance measurement:

25 Commission on Trial Court Performance Standards, *Trial Court Performance Standards with Commentary*, 1997.

26 This description is drawn from Pamela Casey, ‘Defining Optimal Court Performance: The Trial Court Performance Standards’, 1998 *Court Review*, Winter, pp. 24-29.

27 Judicial Council of California, *California Rules of Court*, 2007, <http://www.courtinfo.ca.gov/rules> (accessed 13 April 2007). Originally adopted as Section 30 in 1995, these standards are now incorporated into the newly reorganized *California Rules of Court* as Standard 10.17.

28 A. Aikman, *The Art and Practice of Court Administration*, 2007 p. 237.

1. the number of proposed measures (68) was too great and the measures appeared complex and seemingly without priority (thus in terms of the model, the reform process failed to develop systematically from elite perceptions of what management reforms are desirable (I) to those that are feasible (J) and from thence to actual content (M) and implementation (N));
2. the courts' information systems were not originally designed to produce the data required for the measures, and manual data collection was too labour intensive (constituting a failure of reform in terms of implementation process (N)) ;
3. the economic pressure on budget resources diminished as the economy recovered, and hence the perceived need to invest resources in performance measurement declined (diminished pressure from national and state socio-economic conditions and policies (D));
4. the institutional separation of the judiciary from other branches of government (G), which insulated the courts from pressures to adopt performance measure like some executive branch agencies were compelled to do; and
5. a lack of consistent leadership prevailed on this issue at all levels of the judicial branch (again, a failure to progress from elite perceptions of the desirable (I) to elite perceptions of what management reforms are feasible (J)).

3.1. Court performance measurement reborn

With renewed fiscal pressure on public sector budgets at the turn of the century and a growing perception that judicial branch institutions have not been as successful as other public sector organizations in advocating for budget resources, interest in performance measurement has been renewed. During the intervening years, it is also true that the attention to appropriate performance measurement grew, as indicated by the publication of numerous works on the subject.²⁹ Thus, referring again to the model, judicial branch leaders experienced another round of economic budget pressure (D) (not as severe as in the early 1990s, but enough to remind them of the pain) and a shift in their perceptions of what reforms are feasible (J). The latter was influenced no doubt by increased awareness of the leap in information technology generally (the World Wide Web and Web-based information services had blossomed over that decade) and of improvements in functionality of court information systems made it possible to think of using automated systems for generating key performance reports.

Propelled by increased demand from judicial branch leadership for assistance in promoting 'effective judicial governance and accountability',³⁰ the National Center for State Courts revisited the TCPS in a series of national meetings. Responding to the widely held view that the measures were too many, and taking into account the emerging literature on the balanced scorecard approach to performance measurement, the NCSC developed CourTools, a set of ten performance measures designed to evaluate a small set of key functions of the court. Drawing heavily on the framework originally outlined in the TCPS and its notion of measuring the performance of the court as a whole, CourTools was designed as a 'feasible few' set of measures that were selected on the basis of three criteria: a) correspondence to fundamental court values; b) balanced

29 See for example M. Brown, *Keeping Score: Using the Right Metrics to Drive World-Class Performance*, 1996; R. Chang *et al.*, *Performance Scorecard: Measuring the Right Things in the Real World*, 2000; R. Kaplan *et al.*, 'The Balanced Scorecard: Measures That Drive Performance', 1992 *Harvard Business Review* (January-February), pp. 71-79.

30 Conference of State Court Administrators, *Position Paper on Effective Judicial Governance and Accountability*, 2001. Available at http://cosca.ncsc.dni.us/white_papers.html

perspective on the work of the court; and c) feasibility and sustainability.³¹ Success factors that the measures seek to illuminate include fiscal responsibility, employee satisfaction and engagement, client-customer satisfaction and effectiveness and efficiency of internal processes.

The ten CourTools measures are:

1. *Measure 1 Access and Fairness* (ratings of court users on the court's accessibility, and its treatment of customers in terms of fairness, equality, and respect)
2. *Measure 2 Clearance Rates* (the number of outgoing cases as a percentage of the number of incoming cases)
3. *Measure 3 Time to Disposition* (the percentage of cases disposed within established time frames)
4. *Measure 4 Age of Active Pending Caseload* (the age of active cases pending before the courts, measured as the number of days from filing until the time of measurement)
5. *Measure 5 Trial Date Certainty* (the number of times cases disposed by trial are scheduled for trial)
6. *Measure 6 Reliability and Integrity of Case Files* (the percentage of files that can be retrieved within established time standards and that meet established standards for completeness and accuracy of contents)
7. *Measure 7 Collection of Monetary Penalties* (payments collected and distributed within established timelines, expressed as a percentage of total monetary penalties ordered in specific cases)
8. *Measure 8 Effective Use of Jurors* (measurement of juror yield (the number of citizens who report for jury duty as a percentage of those summoned) and juror utilization (the number of prospective jurors actually used as a percentage of those who reported for jury duty))
9. *Measure 9 Employee Satisfaction* (ratings of court employees assessing the quality of the work environment and relations between staff and management)
10. *Measure 10 Cost per Case* (the average cost of processing a single case, by case type).³²

Citing a) the need 'to promote judicial governance and accountable state Judicial Branch institutions that provide the highest quality service to the public'; b) the need to improve understanding of the judiciary's role to make it less likely that there is interference with 'the judiciary's ability to govern itself'; and c) the need 'to develop benchmarks for court performance measures so judiciaries can assess their own progress and allow courts to compare their performance with similarly situated courts', in August 2005 the Conference of State Court Administrators called upon state courts to implement performance measures and upon the National Center for State Courts to create a national clearinghouse of performance data.³³ What is new here, in comparison to the earlier endorsement of the TCPS, is the explicit call for cross-court comparison and the acknowledgment of both managerial and political reasons for measuring performance. What is also new is the fact that some states and individual courts are actually implementing performance measures. A brief review of these efforts follows below.

31 Brian Ostrom, *Why Measure? Five Reasons to Assess Court Performance*, 2005. Accessed 21 May 2007 at http://www.ncsconline.org/D_Research/CourTools/CourToolsWhitePages-v4.pdf

32 Full definitions and detailed descriptions of these measures are available from the National Center for State Courts at <http://www.courttools.org>

33 Conference of State Court Administrators, *Resolution 14 In Support of Measuring Court Performance*, 2005. <http://cosca.ncsc.dni.us/Resolutions/CourtAdmin/resolutionMeasuringCourtPerformance.html> (accessed 1 April 2007).

3.1.1. State-level performance measurement

The state of Utah has begun to implement the CourTools measures, proceeding measure by measure, statewide. The results of those measurements are published on the state courts' public Web site, www.utcourts.gov, and data for most measures are available at the aggregate statewide level as well as at the local jurisdiction level. The state of North Carolina has developed a Court Performance Measurement System (CPMS) that is also available on its public Web site, reporting data on three measures related to caseload management: clearance rate, time to disposition and age of active pending cases.³⁴ The state of California is currently pilot testing all ten CourTools measures in four courts, with the intention of building reporting capacity on most of the measures into the new statewide California Case Management System. The state of Arizona has major work underway in its largest superior and municipal courts pilot testing many of the CourTools measures. The Arizona judicial branch strategic plan, 'Good to Great: A Strategic Agenda for Arizona's Courts 2005-2010' includes a section on accountability, within which the judicial branch commits to establish performance and operational standards and measures for the courts, based on the CourTools.³⁵ The state of Oregon has adopted a set of budget-related performance measures, which include 17 measures consistent with, identical to, or supplemental to the CourTools measures.³⁶ Finally, the state of Massachusetts has initiated a comprehensive effort to utilize the measures related to caseload management to revitalize that state's court system: clearance rate, time to disposition, age of pending cases, and trial date certainty. Massachusetts has taken the additional step of setting statewide goals for its measures and publishing results.

3.1.2. Court-level performance measurement

Individual courts are also taking up performance measurement, without waiting for broader statewide efforts to get underway. These range from large urban courts like Harris County, Texas, (which includes the city of Houston), Hennepin County, Minnesota, (which includes the city of Minneapolis) and Maricopa County, Arizona, (which includes the city of Phoenix) to small courts of one to six judges in rural areas. Results of the Fourth District Court in Hennepin County and for Maricopa County are available on the court Web sites.³⁷ A good example of a strong effort underway in a midsize urban court is the work being done in Yuma, Arizona, in the superior court, which is reporting results for CourTools measures 1 (Access and Fairness), 9 (employee satisfaction), 8 (effective use of jurors), 3 (time to disposition), 2 (clearance rate) and 10 (cost per case).³⁸ Three smaller rural courts in Indiana, Texas and Ohio have posted the results of their first round of performance measurement, along with management recommendations for actions to take based on the results.³⁹ These courts represent a representative, but not exhaustive sample of the breadth and depth of local-level court performance measurement initiatives across the US.

34 See the Web site of the North Carolina Administrative Office of the Courts at <http://www1.aoc.state.nc.us/cpms/login.do>.

35 Arizona Supreme Court, *Good to Great: A Strategic Agenda for Arizona's Courts 2005-2010*, 2005. <http://www.supreme.state.az.us> (accessed 13 April 2007).

36 A. Phalan, *The Oregon Judicial Department Performance Measurement System*, presentation to the COSCA 2005 mid-year meeting, December 9, 2005, San Diego, California.

37 For Minnesota see www.mncourts.gov/district/4/?page=394 and for Arizona see <http://www.superiorcourt.maricopa.gov/publicinfo/annualrep/FY2006AnnualRpt.pdf>.

38 See <http://www.co.yuma.az.us/courts/dashboard.htm>

39 See the Lubbock, Texas court report at <http://www.co.lubbock.tx.us/DCrt/Reports.htm>; see the Morrow County, Ohio report at <http://morrowcountycpc.com/>; see the Tippecanoe County, Indiana report at www.courttools.org on the Online Community home page.

3.1.3. The perception of feasibility

What accounts for the change in elite (judicial branch leadership) perceptions of the feasibility of performance measurement? One of the key failure points in the previous reform movement based on the TCPS was the inability or unwillingness to move beyond rhetorical statements of the desirable (I) to an actual performance measurement programme (M). Several factors appear to have contributed to this shift.

First, in order to create the perception of feasibility (and ultimately the possibility of comparative evaluation), the CourTools were specifically designed to provide precise data definitions and analytical methodologies to develop results that are credible. State and local court leaders have long known that much of their data is unreliable at worst and unverified at best. Thus, even those who embraced the notion of performance measurement were not prepared to believe the results of any attempt to measure it that relied on current, unrevised data. A necessary prerequisite was thus the development and dissemination of standard definitions of case types (that is, the definition of a contract case, fraud case, tort, etc.) and case counting rules (that is, the relationship among the number of charges/causes of action, the number of defendants/plaintiffs, and the number of cases). This was accomplished through the publication of the *State Court Guide to Statistical Reporting* in 2003, which codifies the universal definitions of case types and standardized counting rules.⁴⁰ For example, in some states *one person* charged with five different criminal offences is counted as *five cases*; the NCSC standard calls for this to be counted as a single case. While states may have local preferences or legal requirements for how they classify or count cases, the guiding principle is to map that data into the standard framework for national comparison purposes.

In addition, detailed specification of the calculation rules, for example, for computing accurate time to disposition, require that case management systems embed the appropriate logic in their queries and reports, and/or that courts utilize appropriate formulas in spreadsheets.⁴¹ Even as simple a notion as the elapsed time from the filing of a case to its disposition involves a series of decision rules and exposes the limitations of current information systems. For example, to count this time correctly, a court must be able to exclude time when the case was removed from court control, *e.g.*, when a civil case is stayed due to the filing of a bankruptcy proceeding in federal court or when a criminal case is halted because the defendant has absconded.

Developing standard definitions, counting rules, and calculations provided the basis for creating a new perception that measurement could be done fairly, accurately, and consistently within and across courts within a given state, and among states. The obvious additional benefit, in the context of diverse state court systems, is that standardizing the precise way the measures are to be taken is the only hope for creating results that can be interpreted and compared in a meaningful way.

3.1.4. Data quality

As noted above, for many if not most courts, the data required either does not exist in usable form or is of dubious quality. This is not surprising, since most courts manage their operations without reference to this data, focusing on the scheduling of events and ad hoc methods of allocation of work among judicial officers and court staff. Although much of this data is reported in the aggregate to the state-level administrative office of the courts, in many states virtually no use is

⁴⁰ National Center for State Courts, *State Court Guide to Statistical Reporting*, 2003.

⁴¹ To ensure error-free calculation, the NCSC has developed a series of spreadsheets for data entry and analysis. See the templates at www.courttools.org on the Online Community home page.

made of it at that level either, other than to publish it in aggregate form in an annual report. At most, a court might receive an inquiry about missing data or wildly different data, but that does not generally lead to any real investigation into the data quality. Of course, for any generalization like this there are exceptions, for example, the states of New Jersey and Minnesota, which have developed and maintained a high level of data quality. This quality is based on the fact that those state offices actually use the data to manage the courts, and court staff and judicial officers have a keen understanding of that fact.

Thus, a critical first step for most courts is the collaborative work among court information technology staff, managers and court staff required to validate and correct their data. Generally this work has proceeded on an iterative basis by producing reports and then investigating anomalies reflected in those reports numerous times until the data is acceptably clean. Several other courts have taken a different path forward. The court management team of the criminal courts of Harris County, Texas, took note of the following:

1. While information technologies come and go, data is eternal, yet little effort and resources are invested in auditing and improving it;
2. Every inquiry begins with the same questions (Do we have the data? Can we get to the data? Is the data readable? Can we convert the data? Is the data reliable?) and the cost of redundant explorations of the answers is significant;
3. The court pays over and over again for slightly different views of the same data, and produces different versions of the truth, depending on exactly how the data is categorized, both of which are expensive and unwise practices.⁴²

This court has undertaken to build its underlying data infrastructure first, before deploying it to produce reports. This stands in contrast to the ‘deploy and fix’ approach to data quality described above. As part of this effort, the Harris County criminal courts subject their data to systematic audit. Errors are identified not only to fix the record, but also to target training to specific staff responsible for the errors. Results for each court are published to the courts, and a healthy competition has evolved among the courts to achieve the lowest error rate. Having initiated a culture that values data quality, the court is now in the final stages of developing Web-based reporting tools and graphical displays of performance measurement results.

3.1.5. Interpreting performance measurement results

Assuming the availability of data relevant to measurement, the second major challenge for court managers is to understand what the results mean. Most court administrators at the state and local levels have managed the courts without any such data for so long that the utility of the information is not at all apparent to them. What this reveals is that some of the fundamental notions of caseflow management, if learned at all, remain invoked at best on an intuitive basis by individual judges or court managers on an ad hoc basis. When the results of a single measure are not understood, it is obvious that the more nuanced use of multiple measures (*i.e.*, the balanced scorecard, and the practice of making conscious management decisions to trade off performance in one dimension (*e.g.*, timeliness) for performance in another (*e.g.*, quality of justice) never appears on the managerial agenda at either the local or state level.

⁴² H. Leverette, *Industrial Strength CourtTools: An Architecture for Court Business Intelligence*, presentation to the Court Solutions Conference, September 18-20, 2006, Baltimore, Maryland. Available online at: <http://www.courtsolutions.org/MS/MS5/page.php?p=451>

Most courts that have engaged in performance measurement in this recent resurgence of interest and implementation have measured key outcomes only once, and cannot yet see the results of any management actions they may have taken in response to those results reflected in measurement at subsequent points in time. In addition, given the as yet small number of courts engaged in performance measurement and the lack of historical data, specific outcomes on measures like clearance rates and collection rates tend to be specific to individual courts and not ripe for comparative analysis. Thus, our discussion below of lessons learned is more of a snapshot of initial insights on some issues arising from two measures that are relevant to understanding the overall effort at court performance measurement. The commentary below is based upon review of performance measurement data from the courts of Arizona, California, Ohio, Texas and Utah.

3.2. Ratings of access and fairness

The first measure in the CourTools set is a measure of access and fairness through ratings of court practice by court users. The survey, administered to all those exiting the courthouse(s) in a jurisdiction on a given day, ask respondents to rank the court on ten items related to access and five items related to procedural fairness if they participated in a hearing before a judicial officer that day. Most courts have discovered that the public rates their services higher than judicial officers and the court's staff and managers might have guessed. No doubt, in part this is due to the skewed impression that the court gets of its customers, based on memory of either the very negative or very positive feedback obtained in person or on voluntary customer feedback survey cards available at the counter. The 'mystery of the middle' – what *most* court users think – has remained a mystery up to now. For most courts, these ratings have indicated a relatively high rating of the court by its constituents. At the same time, courts are gaining valuable information on what parts of their constituency feel less well served. Serving communities with diverse populations is a challenge faced by many urban courts and by some rural courts with immigrant populations who typically make up a significant part of the agricultural and service sector labour force. On the whole, the surveys to date indicate that the courts are, for the most part, perceived as equally accessible to court users regardless of age, sex and race/ethnicity, or type of case that brought them to the court.

The second part of the survey asks five questions related to procedural fairness. Courts are also learning, as the research on procedural justice has consistently shown, that people's evaluation of the court is not based on whether they win or lose, but on whether they were treated with respect and feel that their side of the story was heard by the court and taken into account. This well-established finding remains outside the purview of many judges and attorneys, however, who continue to believe that litigants are focused on the *outcome* ('winning' or 'losing') rather than on the perceived fairness of the process.⁴³ This disjuncture leads to unintended consequences in the courtroom, when judges and attorneys fail to create a process and a language that allows litigants to feel they have had their say and understood their day in court.

Conducting the access and fairness survey is also generating comparative data at the local level. Court managers and judicial officers are contrasting the ratings of the court's various divisions (civil, family, criminal, traffic, juvenile, etc.) and locations. The managerial challenge

43 The research of Thomas Tyler has demonstrated this finding repeatedly. For an introduction to that research, see T. Tyler, 'What is Procedural Justice?: Criteria Used by Citizens to Assess the Fairness of Legal Procedures', 1988 *Law and Society Review* 22, pp.103-135. For a recent confirmation of this disconnect between litigants and judges and attorneys, see Judicial Council of California, *Trust and Confidence in the California Courts, Phase II*, 2006, p.36.

for courts whose management culture is typically reactive, hierarchical, and focused on fixing errors and responding to crises, is to identify areas of excellence and promote courtwide learning and sharing of effective practices, and to do so in a way that involves the court staff.

More generally, it appears true across most courts that while the courts have invested significant resources in the Web sites, the public is largely ignorant of those sites and does not take advantage of functionality (online payment, online traffic school, court calendars, procedural information and forms, etc.) that could make it easier for them to conduct their business, and in some cases avoid a trip to the courthouse in the first place. The public generally finds that it still takes too long to do court business, that court forms are not easy to use or clear, and that the court hours of operation are not convenient.

The relevance of this point is that it illustrates that communicating with their many and varied constituencies is a challenge that is frequently underestimated by courts. Mere publication of information on a Web site or in an annual report does not constitute effective communication. To the extent that courts are measuring performance and demonstrating a commitment to improvement and accountability, this information is not well disseminated. Part of the value of performance measurement lies in the publication of its results, and shaping the public and political discourse about courts. Referring again to the model outlined above, the connection needs to be made between reforms actually achieved (O) and the elite perceptions of what is desirable and possible (I and J), which in turn shape the content of the next iteration of the reform package (M).

3.3. Employee satisfaction

The CourTools measures include a measure of employee satisfaction. This is a survey of approximately 20 questions gleaned from management literature that tap into key dimensions of the work experience. Typically, the survey is administered anonymously as an online survey; response rates in more jurisdictions are approximately 65 %. Reasonable observers of a typical US state courthouse in most jurisdictions would agree that court staff work very hard at their jobs. And many courthouses are far from modern facilities, affording often substandard working conditions and outdated information technologies. What is remarkable is not that court employees work so hard, but that they continue to work hard year after year, without any real feedback about the results of their work. Informal inquiries with court staff reveal that they measure their performance by various ad hoc measures like how high the pile of paper on their desk is on Friday or whether they have had 15 minutes of uninterrupted time at their desk during any day of that week. At best, they might have a sense of some of their outputs – *e.g.*, whether all the notices in a case got mailed out on time.

The employee satisfaction survey provides some insights into the paradox of the hard-working court employee. In analysis of the CourTools survey that categorizes the survey questions into subscales, we find that court employees' values align highly with the mission and vision of the court's higher purpose – providing justice. Court staff believe the work of the court is important and respected in the community, and they experience the connection between the work that they do and the mission of the court. Overall, court employees enjoy their work and are proud to work at the court and it appears that it is these beliefs that motivate their hard work. These findings are all the more impressive given how little information employees get from either their managers or management reports of any kind on the court's performance.

Just as with the surveys of the public, courts administering this survey gain intra-court comparative information that allows sharing of effective managerial practices across locations and divisions. In some instances, the courts have discovered differences in the perceptions of

court workplace management and practices managers and staff, which create fertile ground for productive dialogue within the court. Engaged, committed court employees who are provided with timely information about the court's management agenda are in the best position to provide quality service to those who use the courts.

Again, the challenge for traditional court managers is to engage the staff in the discussion of and solution to the issues identified by court employees. In addition, for some managers, employee satisfaction is seen as irrelevant to their work, while others view it more appropriately as an intermediate outcome that has a direct impact on the outcomes experienced by users of the court (being able to access information, being treated with courtesy and respect, being able to conduct their court business in a reasonable amount of time).

4. Looking to the future

As Pollitt and Bouckaert point out, separating rhetoric from practice and estimating trajectories of reform is risky business, especially since reform packages like performance measurement 'display a considerable rhetorical dimension, playing harmonies on the styles and ideas of the moment'.⁴⁴ And at the local or state level, it is one thing to collect and publish data, it is quite another to use it to make management decisions; this remains the fundamental challenge to effective management of state courts. The vision of this overall project was well summarized by the Massachusetts courts: 'transforming the culture of the Trial Court – a transformation whereby empirical data inform policies and drive management decisions, enabling us to increase our accountability and assess our progress, while maintaining our unwavering commitment to quality substantive justice'.⁴⁵

4.1. Institutionalization

The institutionalization of reform through the administrative system (L in the model) is where many a good idea is created or destroyed. A significant issue for the US state courts is that court administrators typically lack the level of professional management training and education required for leading a complex public institution like the court. Most have come up through the ranks and are operational experts; few have formal education or practical experience in managing complex organizations. On the judicial side of the court management team, most judges are, by inclination and training, not strong managers. An additional weakness of the state court leadership structure is the rotating nature of the chief justice and presiding judge positions, which creates discontinuity within the judiciary as well as in relationships with the other two branches of government.⁴⁶ Creating the educational and professional infrastructure for developing skilled management is a strategic goal whose value is perhaps still underestimated.

Currently, a review of which states and courts are undertaking performance measurement initiatives suggests that for the most part, the key factor in implementing court performance measurement is judicial leadership (elite perceptions (I, J) in the model). It is not the states that are experiencing the worst of the political attacks that are leading this effort, nor is it the states whose judicial branch budgets are most constrained. Rather, it is the relatively well-resourced

44 C. Pollitt & G. Bouckaert, *Public Management Reform: A Comparative Perspective*, 2000, p. 35.

45 Administrative Office of the Trial Court, *Enhancing the Delivery of Quality Justice: Report of the Court Metrics Project, Calendar Year 2006, 2007*.

46 The terms of local court presiding judge (aka chief judge, administrative judge) are typically only two years. See National Center for State Courts, *Key Elements of an Effective Rule of Court on the Role of the Presiding Judge*, 2005. Accessed 22 May 2005 at http://www.ncsconline.org/D_Research/Documents/Res_JudInd_ElementsofaRule_final2.pdf

states and courts with truly visionary leadership that are choosing to reform court management in this way.

4.2. Interpretation

On the analytical side, the challenge of comparative analysis within and across countries of even basic measures looms large. It is obvious from trying to interpret earlier efforts by Dakloias⁴⁷ and her colleagues at the World Bank that data without context, or from wildly different contexts, is difficult if not impossible meaningfully to compare. Certainly this is true in the European context, as part of the process of European integration and the definition of a 'European' justice that is more than the sum of its national parts. The work of Philip Langbroek and Marco Fabri and their colleagues is illustrative in this respect, as these researchers attempt to create meaningful frameworks for comparative analysis of judicial systems.⁴⁸ The work of Zelko Sevic in South Eastern Europe is similarly commendable in this respect.⁴⁹ Through this work, these scholars are filling in a major knowledge gap: how public sector reform efforts in general and performance measurement efforts in particular apply (or not) to judicial systems. As described above, the US state courts suffer from this problem within their own national borders even with a common legal framework; the legacy of their local origins persists in their structures, procedures, data definitions and counting rules, which vary widely. Perhaps some of the strategies for overcoming this variation will be useful for European practitioners and scholars to consider.

Beyond the basic measures discussed here, much remains to be done to define appropriate measures to capture fairness, consistency and quality of justice. In the US, consistency in sentencing is discussed but not evaluated at the local or state level, and the same is true for the quality of written opinions at the trial court level. Beyond these complex issues lies the challenge of the overarching measurement of judicial independence within a state or national framework.

History suggests that advances in judicial governance and accountability through performance measurement will not be undertaken solely or even primarily due to concern for social progress, human rights, or newfound respect for the rule of law. As the economy continues to become increasingly global, the industrialized countries no less than the developing world will compete to attract and retain economic investment. Here perhaps Europe and the US share something else in common: the movement of capital to locations perceived as more hospitable and profitable to business. Judicial systems are one instrument for creating and maintaining a healthy climate for economic growth. Those engaged in the judicial services must ensure that the fundamental vision of fair and impartial justice free from political and economic interference is not lost in the long march to any future model of national or supranational market economy and political democracy.

47 M. Dakloias, 'Court Performance Around the World: A Comparative Perspective', 1999 *The Yale Human Rights & Development Law Journal*, http://www.yale.edu.yhrdlj/vol02/dakloias_maria_article.htm.

48 See for example: P. Langbroek *et al.*, (eds.), *Case Assignment to Courts and Within Courts*, 2004; M. Fabri *et al.*, *The Administration of Justice in Europe: Towards the Development of Quality Standards*, 2003; M. Fabri *et al.*, *L'administration de la justice et l'évaluation de sa qualité*, 2005.

49 Z. Sevic (ed.), *Gauging Success: Performance Measurement in South Eastern Europe*, 2004.