

BUILDING A BETTER COURT

Measuring and Improving
Court Performance and
Judicial Workload in Child
Abuse and Neglect Cases



NATIONAL COUNCIL OF
JUVENILE AND FAMILY COURT JUDGES



The American Bar Association, Center on Children and the Law

The National Center for State Courts

The National Council of Juvenile and Family Court Judges

ACKNOWLEDGMENTS

This project was funded by the David and Lucile Packard Foundation. The Foundation's generous funding has not only allowed our organizations to maintain and expand past efforts to share ideas and strategies, pool resources, and continue to keep each other informed about court reform projects, efforts, and successes throughout the United States, but also funded the development of this *Guide* and companion *Toolkit*. Special thanks to Lucy Salcido Carter, our grant officer, as well as her successor, Sandra Bass, for their continuing support and for helping us get the Packard Foundation involved in this project. We also wish to thank the Office of Juvenile Justice and Delinquency Prevention, Office of Justice Programs, U.S. Department of Justice for providing support through its Child Victims Act Model Courts Project.

Our collaborative team would like to acknowledge the many individuals who contributed to the development of this *Guide* and *Toolkit*, including David Steelman, Denise Dancy, and Lynn Levey from the National Center for State Courts; Melissa Litchfield, Dr. Shirley Dobbin, Chris Bailey (M.A., J.D.), Cheryl Davidek, Dr. Dionne Maxwell, and Jason Oetjen from the National Council of Juvenile and Family Court Judges; and Yvonne Brunot from the American Bar Association (ABA), Center on Children and the Law. Neal Kauder, VisualResearch, Inc., helped with information graphics and supervised report layout and design. Kara Humpheys, of karacreative, was responsible for layout, design, and report production.

Special thanks are owed to all of those individuals who helped arrange each of our project site visits and who greatly facilitated both our work on-site and our access to data. Finally, we would like to express our sincere appreciation to the many judges and court personnel who allowed us into their courts and communities and who gave so generously of their time and expertise.

Support for this research was provided by the David and Lucile Packard Foundation. Points of view and opinions expressed here are those of the authors and do not necessarily represent the views or policies of the National Center for State Courts, National Council for Juvenile and Family Court Judges, the David and Lucile Packard Foundation or individuals who cooperated with the research.

The views expressed herein have not been approved by the House of Delegates or the Board of Governors of the American Bar Association and, accordingly, should not be construed as representing the policy of the American Bar Association.

FOREWORD

The *Guide* and companion *Toolkit*, developed with funding from the David and Lucile Packard Foundation, are designed to help courts handling child abuse and neglect cases make real and sustained advances in improving court performance by successfully measuring court performance and judicial workload needs. By improving their own performance, courts will ultimately improve outcomes for abused and neglected children.

The *Guide* and *Toolkit* are the result of the four-year combined efforts of the American Bar Association (ABA) Center on Children and the Law, the National Center for State Courts, and the National Council of Juvenile and Family Court Judges. Our intent was to synthesize a set of methods from court performance measurement and workload assessment site visits that would provide tools for the assessment of performance; provide tools for evaluating judicial workload; and provide tools that can be tailored to individual courts and their measurement needs and data collection capabilities. Ultimately we hope that courts will use these methods and accompanying instruments not only to assess and improve performance but also to build their internal capacity for self-examination and self-improvement.

This *Guide* and *Toolkit* outline a process that courts can use to measure court performance and judicial workload. They have been designed to be meaningful to the largest number of courts as possible, regardless of individual differences in court structures and procedures, and to represent our collective knowledge and experience about performance measurement and judicial workload

assessment – knowledge and experience that was greatly informed by conducting numerous site visit studies as part of this Packard Foundation-funded project. We present the lessons we have learned during instrument development, data collection site visits, and numerous collaborative meetings to refine our procedures and data elements. The tools are not only complementary to the federal requirements and Child and Family Service Reviews (CFSRs) but also useful as a gauge of progress toward aspirational or “best practice” goals. While designed to be meaningful for most courts, the tools are readily adaptable to an individual jurisdiction’s measurement needs and data collection capabilities. As a result of using the *Guide* and *Toolkit*, courts should have not only reliable baseline measures of performance and workload, but also a process in place to continue their assessment efforts beyond initial evaluations.

To obtain a copy of the companion *Packard Court Performance Measurement and Judicial Workload Assessment Toolkit*, and for additional information about accessing technical assistance to implement the procedures outlined in this *Guide*, please contact one of the partner organizations.

American Bar Association, Center on Children and the Law
Mark Hardin, Esq., Director, Child Welfare (202) 662-1750

National Center for State Courts, Research Division
Victor E. Flango, Ph.D., Vice President (757) 253-2000

National Council of Juvenile and Family Court Judges,
Permanency Planning for Children Dept.
Mary V. Mentaberry, Director (775) 327-5300

By improving their own performance, courts will ultimately improve outcomes for abused and neglected children.

TABLE OF CONTENTS

Chapter 1

Introduction to Improving Court Performance in Child Abuse and Neglect Proceedings	1
Background	1
The Need for Performance Measurement	3
The Need for Judicial Workload Assessment	6
The Packard Guide and Toolkit	8

Chapter 2

The Court Performance Measures	12
What Are the Desired Outcomes for Children and Families?	12
How Should Outcome Measures for Courts Differ from Outcome Measures for Agencies?	14
Can Courts and Agencies Actually Provide These Measures?	20
Deciding Which Performance Measures to Use	21
Methods for Obtaining Court Performance Measures – The Importance of Triangulation	21
Quantitative and Qualitative Court Performance Measurement Methods – A Brief Outline of Strengths and Limitations ...	22
Conclusion	27

Chapter 3

Assessing Judicial Workload	28
Before the Workload Assessment Begins	29
The Workload Assessment: Step-by-Step	31
An Independent Perspective Is Necessary	45
The Delphi Approach	46
Conclusions	51

Chapter 4

The Planning Process – How Your Court Can Make It Happen	53
Introduction	53
Phase 1: Planning	53
Phase 2: Implementation	56
Phase 3: Wrap-Up	57
Lessons Learned	58
Phase 1: Planning Lessons	59
Phase 2: Implementation Lessons	62
Phase 3: Wrap-Up Lessons	70

Chapter 1

Introduction to Improving Court Performance in Child Abuse and Neglect Proceedings

BACKGROUND

Approximately three million children are reported as victims of abuse and neglect by parents or other caregivers each year.¹ These allegations of maltreatment are investigated by the appropriate child welfare agency, and approximately one-third of reports are substantiated.² Many of the children found to be maltreated enter the foster care system, and approximately 15 percent of the victims require court action of some sort.



Child Maltreatment, 2000
Source: U.S. Department of Health and Human Services

Today, more than ever before, courts are playing an essential role in ensuring the safety, permanency, and well-being of abused and neglected children. Courts play

a key role in determining whether children will be removed from their homes, how long they will remain in foster care, and where they will permanently reside. It is absolutely essential that we improve court performance in child abuse and neglect proceedings – perhaps more so than in any other area of government. Given the number of cases and the high stakes for children and families involved, every effort must be made to measure performance, identify areas in need of improvement, chart progress, and provide the stimulus to improve society’s response to child maltreatment.

In recent years, the federal government has concentrated with increasing focus on achieving safety, permanency, and well-being for abused and neglected children. Through a combination of legislation, regulations, and executive policy guidance, the federal government has encouraged agencies, courts, and other stakeholders to work together to achieve safe, permanent, and loving homes for children involved in the child welfare system. Courts are holding more hearings, addressing more issues in each hearing, dealing with more participants in court, and meeting new and stricter case-processing deadlines. Much of the added responsibility and accountability stems from federal child welfare reform laws, including the Adoption and Safe Families Act of 1997 (ASFA).

Building on the mandates of ASFA, the federal government is working with state child welfare agencies to assess state performance through the Child and Family Services Reviews (CFSRs), a process that examines child welfare outcomes and state systemic factors – both of which include significant legal and judicial dimensions. Each review produces a final report describing a state’s

¹ U.S. Department of Health and Human Services, National Center on Child Abuse and Neglect, *Child Maltreatment 2000* (Washington, DC: Administration for Children and Families 2002).

² Half of investigated allegations are unsubstantiated, 4 percent are found to be false reports, and the remaining investigations are closed by findings of “unknown” or “other.”

Excessive delays, rushed court hearings, lack of adequate or timely notice, brief or inaccurate judicial findings, and persistent lack of court and agency collaboration continue to be systemic problems.

performance with respect to safety, permanency, due process, and timeliness, and typically includes specific assessments of court performance. Based on the CFSR final report, the state then develops a Program Improvement Plan (PIP) designed to achieve systemic improvements and better outcomes for children involved in the child welfare system.

The federal government is also continuing to work with state courts to improve their handling of child welfare cases through the Court Improvement Program (CIP). With the help of federal grants provided through the CIP, courts are required to assess their own performance and to develop and implement plans for improvement. Federal legislation authorizing the CIP directs courts to work with their state child welfare agency to help implement the state PIP. Federal legislation, regulations, and policy issuances for both the CIP and CFSRs recognize that *courts*, as well as state child welfare agencies, are crucial stakeholders in achieving positive outcomes for abused and neglected children involved in the child welfare system. Court performance impacts the combined performance of courts and child welfare agencies in achieving safety and timely permanency for children.

Federal initiatives have helped to dramatically increase interest in improving the performance of courts handling abuse and neglect cases around the country. In recent years, with the help of the federal government, many states have enacted improved legislation governing the court process, refined court procedures, and provided multidisciplinary training for court improvement. In addition, increasing numbers of courts have realized the need to enhance their relationships with state child

welfare agencies in order to effectively handle their abuse and neglect court cases. Despite these advances, however, many courts are not yet able to achieve excellence in handling child welfare cases. Excessive delays, rushed court hearings, lack of adequate or timely notice, brief or inaccurate judicial findings, and persistent lack of court and agency collaboration continue to be systemic problems. In times of reduced resources, increasing federal pressure, and heightened accountability, courts need to reflect on how they do the work that they do, how to better manage their calendars, how to improve the productivity of their staff, and ultimately how to ensure better outcomes for children and families.

Yet, few courts are able to effectively measure their performance in child abuse and neglect cases. Moreover, the concept of performance measurement is still a relatively new concept for juvenile and family courts. But, without the ability to measure performance, establish benchmarks, and track progress, courts will never be able to fully achieve needed systems' reform and improved outcomes for children.

Two types of measurement activities are vital to court improvement: (1) performance measurement to establish baseline performance, identify areas for reform, and chart their own progress in meeting deadlines and other goals, and (2) judicial workload measurement in order to track the resources courts have and persuasively argue for what they need to make major gains in their performance. To achieve long-term court improvement, courts must have the capacity to engage in ongoing performance measurement and judicial workload assessments.

The court performance measures and strategies presented in this *Guide* and accompanying *Toolkit* reinforce current federal reform initiatives, including the CFSRs and the CIP. They strengthen the CFSRs and state Program Improvement Plans (PIPs) by helping states identify areas of court reform that are essential to achieving better outcomes for children. They assist the CIP by providing courts with assessment tools that enable courts to engage in a systematic process of court performance measurement and judicial workload assessment.

The *Guide* and *Toolkit* are designed to help courts handling abuse and neglect cases make real and sustained advances in improving court performance by successfully measuring court performance and judicial workload needs. By improving their own performance courts will ultimately improve outcomes for abused and neglected children.

THE NEED FOR PERFORMANCE MEASUREMENT

The public holds both courts and child welfare agencies accountable for outcomes for abused and neglected children. Ultimately, legislatures or other funding bodies that provide resources may impose performance measures to ensure accountability if courts and agencies cannot agree on a mechanism for assessing their own performance. Even beyond basic requirements of accountability, however, courts and child welfare agencies should always be asking: “How can we most effectively meet the needs of children and their families?” and “Are families receiving the services they need and are services delivered in a fashion to produce the desired results?”

Performance measures are the yardstick by which courts and child welfare agencies can measure the success they are achieving and the progress they are making.

In order for courts to improve and sustain system improvements, the development of objective measurements of practice is critical. Courts, like child welfare agencies, must not only focus on timeliness of case processing and decision making, but also on the quality of the process and related outcomes. For example, courts make key safety decisions – two of the most important being deciding whether to remove a child from the home and deciding whether to return a child to the home. Measuring the appropriateness of these decisions, and the impact of these decisions on children, motivates courts to critically examine their decision-making process (e.g., the quality and comprehensiveness of the information upon which the decision is based), the outcomes and impacts of those decisions, and how to enhance future decision making. Courts must focus on ensuring secure permanent homes for children in foster care, and must determine their effectiveness in achieving permanency. Courts also need to determine how well they are protecting the rights of those children and adults who come before them. Finally, courts need to set *aspirational* performance goals in each of these areas – goals that are designed to focus their efforts, motivate their staff, and evaluate their achievements.

Performance measures are the yardstick by which courts and child welfare agencies can measure the success they are achieving and the progress they are making.

The critical importance of measuring performance*

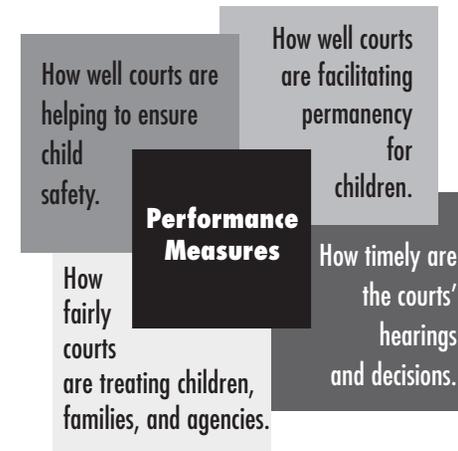
- What gets measured gets done
- If you don't measure results, you can't tell success from failure
- If you can't see success, you can't reward it
- If you can't reward success, you're probably rewarding failure
- If you can't see success, you can't learn from it
- If you can't recognize failure, you can't correct it
- If you can demonstrate results, you can win public support

From D. Osborne and T. Gaebler (1992). *Reinventing Government: How the Entrepreneurial Spirit Is Transforming the Public Sector*. Reading, MA: Addison-Wesley Publishing (pp. 146-152).

Performance measurement identifies those courts that are achieving positive results. This helps to identify and document good practice, while informing the field.

At present, few courts are able to effectively measure their performance in abuse and neglect cases. And while for-profit businesses have long taken for granted the need for performance measurement, it is still a relatively new concept for our nation's courts. Unless courts handling abuse and neglect cases have this essential information they cannot and will not know what types of improvements are most effective and whether their attempts at improvement are working – for the court, the broader child welfare community, and for children and families.

Performance measurement helps courts diagnose and assess what they need to improve and to periodically reassess areas needing improvement. Initially, courts need to establish a *baseline* of current practice from which they can then build improvement efforts, followed by regular practice *reassessments* as reforms are put into place, tried, and retooled. Performance measurement can be described in the following diagram.



Court performance measurement is important to the successful implementation of the Adoption and Safe Families Act (ASFA). ASFA makes the safety and well-being of abused and neglected children its paramount goal. As explained above, court performance measurement can help courts improve their record in safeguarding abused and neglected children. Court performance measurement enables courts to measure whether they:

- Comply with ASFA's timelines;
- Hold substantive, meaningful, and timely hearings;
- Achieve timely permanency for children with minimal disruption in placement;
- Conduct permanency hearings that decide on permanent outcomes for children – such as return home, adoption, legal guardianship, and permanent placements with relatives; and
- Provide procedural protections for parties – such as notice to parents and foster parents.

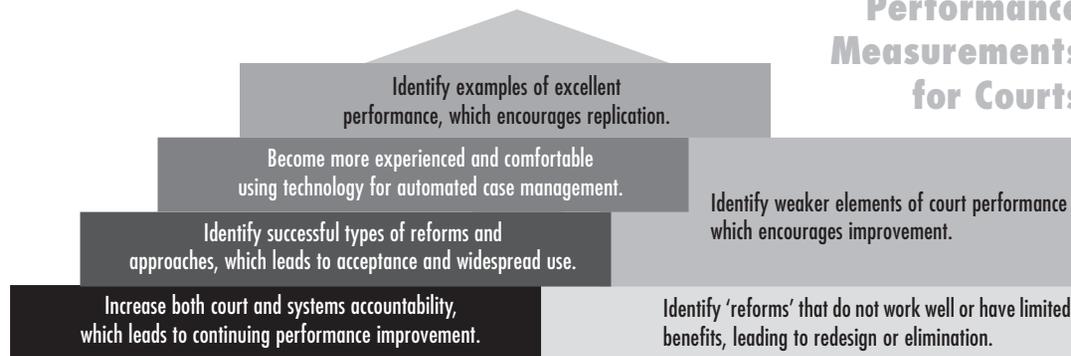
Performance measurement identifies those courts that are achieving positive results. This helps to identify and document good practice, while informing the field. For example, by measuring judicial timeliness in various courts, court administrators can identify those courts with the most timely judicial processes. After identifying such courts, court administrators can further examine how these courts accomplish such timely decisions, while ensuring the quality of the decision making, and what other courts need to do to replicate such results. By establishing concrete measures of judicial activities, states will be able to set certain goals for the courts.

Court performance measurement is needed to demonstrate the value of successful pilot or demonstration projects. Where successful pilot projects lead to more timely permanency, measurement of the result can help courts prove that such reforms are cost-effective and should be replicated. For example, courts that have the capacity to link reforms to shortened lengths of stay for children in foster care can prove that carefully targeted resources can actually reduce the overall level of public expenditures for the child welfare system.

Performance measurement is also critical to the long-term expansion and sustainability of successful reforms. Only with a clear demonstration of success will most courts be able to access the funding needed to successfully develop, and sustain, reforms. In the current fiscal environment, funding for improvements is possible only with compelling and objective proof that improvements will achieve concrete and favorable results for families and children.

In the long term, performance measurement will help courts for reasons stated in the diagram below. Ultimately, measuring court performance will help bring about improved outcomes for children, helping to achieve those outcomes measured by CFSRs and intended by ASFA.

Long-Term Performance Measurements for Courts



Along with the rapidly developing and changing nature of abuse and neglect proceedings, courts have fallen behind in adjusting to newly expanded workload requirements.

THE NEED FOR JUDICIAL WORKLOAD ASSESSMENT

Equally critical to court improvement is objectively assessing judicial workload for courts handling abuse and neglect cases. While the availability of sufficient resources does not guarantee good performance or positive outcomes for children, the lack of adequate resources will almost always hamper a court's performance. Judicial workload improvements are a key component of improving court performance.

Insufficient judicial time and/or resources results in foreseeable consequences for courts:

- Increased time between scheduling contested hearings due to decreased amounts of docket time
- Delayed contested hearings due to insufficient docket time
- Insufficient time during hearings to focus on the safety, permanency, health, and well-being needs of the individual child
- Lack of time to adequately prepare for and conduct hearings, which may result in more errors and impact safety and timely permanency for children
- Inadequate time for off-the-bench judicial activities resulting in a lack of necessary collaboration with child welfare and other agencies and service providers

The specific and individualized character of abuse and neglect cases affects judicial workload needs for this case type. Along with the rapidly developing and changing nature of abuse and neglect proceedings, courts have fallen behind in adjusting to newly expanded workload requirements. Child welfare litigation requires an updated

approach to workload assessment – an approach that takes the following into account:

- The new types of hearings now required by federal and state law
- The increased frequency of certain types of these new hearings, whether imposed by law or recommended by national standards
- The growing number of issues that state laws and national standards now require judges to address in each type of hearing (e.g., agency efforts to locate missing parties, involvement of fathers, reasonable efforts to prevent placement, child support, timely assessments and services, and many others)
- The increasing number of attorneys, parties, and participants in abuse and neglect cases (e.g., putative fathers, noncustodial parents, relatives, foster parents, attorneys for noncustodial and putative parents, treatment providers, and others)
- The increasing procedural protections for parties (e.g., mandatory judicial findings and notice requirements)
- The additional, stricter time lines for hearings, decisions, and petitions

Each of these increased demands requires additional judicial time per case. Increased work with key groups in the community increases the off-the-bench responsibilities of judges hearing a child abuse and neglect docket. The growing complexity of cases requires increased judicial interaction with the child welfare agency and the involvement of additional agencies and organizations in abuse and neglect cases. Judges handling abuse and neglect cases now frequently meet with key agency representatives to seek their cooperation in developing a more effective and efficient court process. The increased complexity of the litigation also requires

judges to focus on their own ongoing training and education, as well as on the training and oversight of attorneys and agency staff performance. Thus, for abuse and neglect cases, judicial workload assessment not only must describe current judicial workloads, but also must determine what additional time and resources are needed to enable the courts to perform at a level of sustainable excellence – in both on-the-bench and off-the-bench responsibilities.³

Calculation of judicial workloads in abuse and neglect cases needs to be objective and reliable in order to convince court administrators and legislative funders to make adjustments to judicial workloads. Sound and objective analysis of judicial workload requirements will help courts understand what resources they need and how to strengthen their requests for those resources. Through workload assessment that is realistic and convincing, which incorporates on-the-bench and off-the-bench responsibilities, courts can provide funders with critical data demonstrating need.

Improved judicial workload assessment can help achieve ASFA's goals in a variety of ways:

- When judges and other decision makers have sufficient time to hear the evidence, ask questions, require parties to produce additional evidence, and make thoughtful decisions, they are better equipped to make individualized decisions that will help ensure children's safety, permanency, and well-being
- When court dockets are not overcrowded courts can initiate and complete hearings within ASFA's timelines

- When judges have the time to carefully review agency reports and testimony and to ask incisive questions, they can help ensure that foster care drift is no longer the path of least resistance for advocates and workers in the system
- When courts have manageable workloads, they can create effective case review systems, provide procedural protections, and make necessary and detailed findings

Judicial workload assessment is also important to the success of state CFSRs. Without manageable judicial workloads key reforms will remain difficult or impossible to achieve. Better workload assessment will contribute to the argument for sufficient case time so that each individual judge has adequate calendar time to devote to each abused and neglected child, improve their decisions, schedule more timely hearings, make more timely decisions, and enter more complete findings. State PIPs should also call for strengthened judicial workload assessment whenever excessive judicial workloads appear to be a barrier to achieving positive outcomes for children. Similarly, state CIP projects should include improved workload assessment.⁴ Improving judicial workload supports court reform efforts aimed at achieving positive outcomes for children and families.

This *Guide* and companion *Toolkit* present an approach to judicial workload assessment designed to help courts (1) establish the amount of time judicial officers *currently* devote to all aspects of abuse and neglect cases (including on-the-bench and off-the-bench activities) and (2) empirically estimate the amount of judicial time needed to fulfill national best practice standards.⁵

³ For further discussion of the need to expand the concept of judicial workload to include off-the-bench activities, see S. A. Dobbin and S. I. Gatowski, "Judicial Workload Estimates: Redefining the Concept of Judicial Work." *Technical Assistance Bulletin* 5, no. 1 (2001), published by the National Council of Juvenile and Family Court Judges, Reno, Nevada.

⁴ The original federal CIP legislation and guidance called for analysis and improvement of judicial workloads in child welfare cases.

⁵ The American Bar Association (ABA), the National Council of Juvenile and Family Court Judges (NCJFCJ), and the Conference of Chief Justices have endorsed a set of judicial best practice standards for abuse and neglect cases known as the *Resource Guidelines*. The standards encourage judges to play a more active role overseeing abuse and neglect cases and describe best practices for conducting child abuse and neglect hearings. See *Resource Guidelines: Improving Court Practice in Child Abuse and Neglect Cases* (Reno, NV: National Council of Juvenile and Family Court Judges, 1995). See also *Adoption and Permanency Guidelines: Improving Court Practice in Child Abuse and Neglect Cases* (Reno, NV: National Council of Juvenile and Family Court Judges, 2000).

THE PACKARD GUIDE AND TOOLKIT

With the generous support of the David and Lucile Packard Foundation, the American Bar Association Center on Children and the Law (ABA), the National Center for State Courts (NCSC), and the National Council of Juvenile and Family Court Judges (NCJFCJ) — three of the nation’s largest and most influential judicial and legal organizations focusing on child abuse and neglect issues — developed this *Guide* and *Toolkit* to help courts improve their performance by addressing the two most critical and challenging areas of court reform:

- Court Performance Measurement – by increasing court accountability through enhanced performance measurement
- Judicial Workload – by assessing judicial workloads to ensure that judges have enough time to make timely, thorough, and well-considered decisions for children and families

The *Guide* and *Toolkit* have been developed to help courts measure and improve their performance, and as part of that performance measurement and improvement, assess their judicial workload needs. The *Guide* and *Toolkit* explain the need for these two areas of reform, set forth methods to address them, provide specific instruments to assist courts in measuring these areas, and explain how to use these instruments. The *Guide* and *Toolkit* present a set of performance measures and a system of measurement comparable to those used by state agencies in CFSRs – so that, like agencies, courts can measure their performance and track their own progress in improving safety, permanency, and timeliness for the children who come before them.

The *Guide* and *Toolkit* have been developed to help courts measure and improve their performance...

The court performance measures cover four basic outcomes:

- Safety – to ensure that children are safe from abuse while under court jurisdiction
- Permanency – to ensure children have permanency and stability in their living situations
- Due Process – to ensure cases are decided impartially and thoroughly, based on evidence brought before the court
- Timeliness – to expedite permanency by minimizing the time from the filing of the petition or shelter care order to the achievement of permanency

None of the performance measures suggest a *standard* or *benchmark* of performance. This is deliberate. The measures were designed to help courts *improve* services to maltreated children and their families. It is therefore important for all courts to accurately measure the level of services currently provided as a benchmark — a local level of performance against which progress can be measured. Courts and child welfare agencies can then establish minimum acceptable standards of performance that should be attained by all courts. They can also set *aspirational* standards that challenge courts and agencies to improve even further. Courts can set aspirational goals that complement and reinforce the state’s PIP goals.

As illustrated on the following three pages, the Packard court performance measures reinforce the AFCARS performance measures and the CFSR standards. Since courts should be expected to collaborate with agencies in applying these measures, our organizations believe it is logical to take into account the performance outcomes developed for child welfare agencies, and designed the Packard court performance measures to be compatible with the AFCARS and CFSR measures.

FEDERAL CFSR

Children are, first and foremost, protected from abuse and neglect.

- Risk of harm to child.
- Recurrence of maltreatment.
- Incidence of child abuse and/or neglect in foster care.
- Children are safely maintained in their own homes whenever possible and appropriate.

PACKARD BUILDING A BETTER COURT

To ensure children are safe from abuse and neglect while under court jurisdiction.

- Percentage of children who do not have a subsequent petition of maltreatment filed in court after the initial petition is filed.
- Percentage of children who are the subject of additional allegations of maltreatment within 12 months after the original petition was closed.

ASFA

Children are, first and foremost, protected from abuse and neglect.

No child should be subject to maltreatment while in placement.

Children are safely maintained in their homes whenever possible and appropriate.

Children have permanency and stability in their living situations & the continuity of family relationships and connections is preserved for children.

- Stability of foster care placement.
- Permanency goal for child.
- Length of time to achieve reunification.
- Length of time to achieve adoption.
- Permanency goal of other planned living arrangement.
- Incidence of foster care reentries.

To ensure children have permanency and stability in their living situations.⁶

Percentage of children who are:

- Transferred among one, two, three, or more placements while under court jurisdiction (this measure should distinguish placements in and out of the child's own home from multiple placements in different environments).
- Percentage of children who reach legal permanency (by adoption, reunification, guardianship or planned permanent living arrangement and other legal categories that correspond with ASFA) within 6, 12, 18, and 24 months from removal (timelines should be adapted to court's specific timelines).
- Percentage of children who do not achieve permanency in the foster care system (e.g., court jurisdiction ends because the child reaches the age of majority).
- Percentage of children who reenter foster care pursuant to court order within 12 and 24 months of being reunified.
- Percentage of children who return to foster care pursuant to court order within 12 and 24 months of being adopted or placed with an individual or couple who are permanent guardians.⁷

Children have permanency and stability in their living situations.⁸

The continuity of family relationships and connections is preserved for children.

⁶ Group consensus was that all measures under "Permanency" should measure stability as well, since the federal child and family services reviews (CFSR) include "stability" as part of overall permanency.

⁷ Since this measure tracks disruption of intended permanent placement to some extent, we recommend it be an "aspirational goal" for courts.

⁸ The W. K. Kellogg Foundation's Families for Kids program, among other organizations, adds a time dimension to the permanency goal – placement in nurturing, permanent homes within one year. This elapsed time goal will be considered here as an integral part of the measure of permanency because it is a shared goal of courts and social service agencies. The timeliness of court processing, however, will be considered part of Goal 4, discussed below.

Process for foster parents, pre-adoptive parents, and relative caregivers of children in foster care to be notified of, and have an opportunity to be heard in, any review or hearing held with respect to the child.

To decide cases impartially and thoroughly, based on evidence brought before the court.⁹

- Percentage of cases in which courts effectuate service of process for both parents prior to adjudication.
- Percentage of cases with documentation that notice is given to parties in advance of the next hearing.¹⁰
- Percentage of children receiving legal counsel, guardians ad litem or CASAs in advance of the preliminary protective hearing.
- Percentage of cases where counsel for parents are appointed in advance of the preliminary protective hearing (0-5 days; 6-10 days; more than 10 days).
- Percentage of cases in which legal counsel for parents and children changes (as well as number of changes in counsel).
- Percentage of cases where legal counsel for parents, children, and government are present at each hearing.
- Percentage of children for whom all hearings are heard by one judicial officer (as well as two, three, or more judicial officers if that information is available).¹¹

Timely service of process for both parents.

Timely provision of notice to all parties in advance of hearings.

Adequate representation and timely appointment of counsel for all parties.

Continuity of counsel and judicial officer.

Sufficient time allotment for hearings.

A court culture that emphasizes follow-up on permanency planning.

⁹ Due process requires notification of involved parties; adequate representation of parties by counsel; an adequate amount of time allotted for hearings; and court emphasis and follow-up on permanency planning.

¹⁰ For most courts this also may be an "aspirational goal" reflecting best practices.

¹¹ To measure the stability of judicial involvement, the principle at work is consistency of decisions and information as well as the avoidance of loss of relationships.

(Under Statewide Information System/Case Review System.)

- Process for periodic review at least once every six months, by court or administrative review.
- Deadline for filing termination of parental rights petitions in accordance with ASFA.
- Permanency hearings within 12 months after a child is considered to have entered foster care and at least once every 12 months thereafter.

To enhance expedition to permanency and timeliness by minimizing the time from the filing of the petition or protective custody order to permanency.¹²

- Average (median) time from filing of the original petition to disposition.
- Percentage of cases that receive a disposition within 10, 30, 60 days after the adjudication of abuse or neglect.
- Average (median) time from filing of the original petition to adjudication.
- Percentage of cases that are adjudicated within 30, 60, 90 days after the filing of the petition.
- Median time from filing of the original petition to finalized termination of parental rights.
- Percentage of cases for which the termination petition is filed within 3, 6, 12, 15, 18 months after the disposition.
- Average (median) time from filing of the original petition to permanent placement.
- Percentage of cases for which a permanency (or adoption) petition is filed within 1, 3, 6 months after the termination order.
- Percentage of cases for which the adoption is finalized within 1, 3, 6, 12 months after the adoption petition.
- Percentage of hearings (by hearing type) not completed within times set forth in statute or court rules (and where available, the reason[s] for noncompletion).

Enhance expedition to permanency and timeliness by minimizing the time from the filing of the petition or protective custody order to permanency.

Average (median) time from filing of the original petition to adjudication, disposition, and permanency that meets or betters state and federal statutory requirements for case-processing time frames.

¹² Two appellate measures are usually included as part of the timeliness goal: (1) percentage of adjudication, disposition, termination, and other judicial decisions that are appealed and percentage overturned on appeal and (2) percentage of cases in which the results of the appeal are received within 1, 3, 6, and 12 months from the date the appeal was filed. The goals are very important and relevant, but cannot be obtained from trial court case files. A separate effort to gather such data from courts of appeal is necessary, but beyond the scope of the measurement strategy outlined in this *Guide and Toolkit*. Note: Interview evidence suggests that many appeals go from the trial court to the intermediate appellate court on an expedited schedule.

¹³ *Supra*, note 5.

Until now, courts have limited their measurement of performance to the timeliness of their proceedings. While the *Guide* and *Toolkit* already encompass much more than the timeliness of proceedings, the ongoing collaborative work of the partnering organizations is aimed at refining and developing better and more comprehensive court performance measures. Use of the *Guide* and *Toolkit*'s performance measurement strategy and instruments in different jurisdictions with different measurement, caseload, and resource challenges will serve to refine the court performance measures and judicial workload assessment process.

Ultimately, the *Guide* and *Toolkit* strive to help courts to measure:

- Their success in helping achieve child safety
- Their success in helping attain permanency for abused and neglected children
- Their procedural fairness toward children, families, and agencies
- The timeliness of their hearings and decisions more comprehensively
- (Eventually) their role in achieving well-being for the children they serve

Chapter 2

The Court Performance Measures

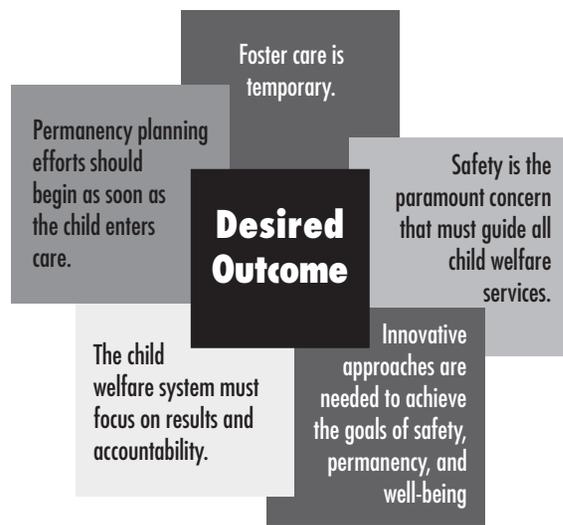
The Packard court performance measures were developed as a result of collaboration by the American Bar Association (ABA) Center for Children and the Law, the National Center for State Courts (NCSC), and the National Council of Juvenile and Family Court Judges (NCJFCJ). Because courts share responsibility with child welfare agencies for the results of interventions into the lives of children and families, the core Packard court performance measures adapt the federal child welfare outcome goals for use by courts handling child abuse and neglect cases – covering the areas of safety, permanency, due process, and timeliness. Our organizations also recognized that the NCJFCJ's *Resource Guidelines* and *Adoption and Permanency Guidelines*¹³ offer excellent aspirational court performance goals for dependency and family courts, and we have incorporated these goals into our methodology.

The authors consider the performance measures outlined in this chapter to be a fundamental and essential core set of measures. While we expect courts to consider additional data of local interest, we strongly urge that courts seriously consider adapting the limited core set of measures presented herein.

WHAT ARE THE DESIRED OUTCOMES FOR CHILDREN AND FAMILIES?

The Adoption and Safe Families Act of 1997 (ASFA) was in part “a response to the fact that more children

were entering the foster care system than were exiting.”¹⁴ Five principles underlie the ASFA and apply to professionals working with families through public and private agencies as well as state courts.^{15,16} These principles are shown in the diagram:



Another important provision of ASFA is the requirement that the Department of Health and Human Services (DHHS) identify a national set of outcome measures that can be used to gauge state and national progress in reaching these goals. A focus on outcomes is also at the heart of the new Child and Family Services Reviews (CFSR) to assess the performance of state child welfare systems.¹⁷ As a result, states must now demonstrate that their programs have actually made a difference in children’s lives.

DHHS published the final list of child welfare outcome measures in the *Federal Register* on August 20, 1999. The outcomes should either reduce or increase specific outcomes as shown:¹⁸

DHHS List of Final Child Welfare Outcome Measures

REDUCE

- Recurrence of child abuse and/or neglect
- Incidence of child abuse and/or neglect in foster care
- Time in foster care to reunification without increasing re-entry
- Time in foster care to adoption
- Placements of young children in group homes or institutions

INCREASE

- Permanency for children in foster care
- Placement stability

Finally, and perhaps most importantly, ASFA clearly and unequivocally establishes three national goals for children in foster care: **safety, permanency, and well-being**. Each of these will be discussed in this chapter and explicitly linked to measures of court performance.

¹⁴ U.S. General Accounting Office, *Juvenile Courts: Reforms Aim to Better Serve Maltreated Children* (Washington, DC: U.S. General Accounting Office, 1999), p. 8.

¹⁵ *Adoption and Safe Families Act of 1997* (P.L. 105-89) signed into law November 19, 1997, amending titles IV-B and IV-E of the Social Security Act. Although these principles are directed at child welfare agencies, courts are legally bound to apply them.

¹⁶ U.S. Department of Health and Human Services, Children’s Bureau, *Guidelines for Public Policy and State Legislation Governing Permanence for Children* (Washington, DC: U.S. Department of Health and Human Services, Children’s Bureau, 1999), pp. 1-5—1-6 (hereinafter referred to as *Guidelines for Public Policy*).

¹⁷ The Child and Family Services review process was published in the *Federal Register* (65 FR 4020-4093) on January 25, 2000.

¹⁸ ASFA requires that these outcomes be reported in an annual report to the Congress. For more information about the Annual Report outcomes see www.acf.dhhs.gov/programs/cb/special/frcwoutc.pdf (an Adobe Reader is required to view this document).

¹⁹ Bureau of Justice Assistance, *Trial Court Performance Standards with Commentary* (Washington, DC: Bureau of Justice Assistance, 1997) (NCJ No. 161570).

²⁰ The measures proposed here are from the Attaining Permanency for Abused and Neglected Children Project, conducted jointly by the National Center for State Courts (NCSC), the American Bar Association (ABA) Center on Children and the Law, and the National Council of Juvenile and Family Court Judges (NCJFCJ), with funding by the David and Lucile Packard Foundation. A partial history of the development of these measures is contained in Victor E. Flango, "Measuring Progress in Improving Court Processing of Child Abuse and Neglect Cases," *Family Court Review* 39 (April 2001): 158-169.

HOW SHOULD OUTCOME MEASURES FOR COURTS DIFFER FROM OUTCOME MEASURES FOR AGENCIES?

While outcome measures for courts are less well defined than measures for agencies, established principles of court performance can be applied to abuse and neglect cases. For example, the National Center for State Courts (NCSC) Trial Court Performance Standards and accompanying measurement system are the culmination of many years of work to develop outcome-based performance standards for courts of general jurisdiction.¹⁹ These Standards have been endorsed by the Conference of Chief Justices, the Conference of State Court Administrators, the American Judges Association, and the National Association for Court Management.

The Trial Court Performance Standards touch upon five fundamental purposes of courts, and these all apply to dependency or child protection court practice as well: 1) Access to Justice; 2) Expedition and Timeliness; 3) Equality, Fairness, and Integrity; 4) Independence and Accountability; and 5) Public Trust and Confidence. These types of outcome measures have broadened the discourse about performance from an emphasis on process alone, such as focusing primarily on how to move cases through the court at a faster pace, to a concern about *outcomes*. Discussion, once limited to court organization and management, now also focuses on how to better serve the public.

Although the Trial Court Performance Standards provide a context to help courts assess their own performance, more is needed to develop and promote standards tailored specifically to child abuse and neglect cases.

Performance measures should give courts feedback on how well they carry out the intentions of the child abuse and neglect laws, how well they protect abused and neglected children from further harm, and how well they help children to grow up in permanent, functional families. Clearly, these goals are just as important to courts as they are to child welfare agencies. On a more systemic level, court performance standards should help courts identify where problem areas exist, and thus where to focus improvement efforts. Indeed, the availability of court performance standards should allow court managers to identify which courts are performing well. National court reform organizations can use these exemplary courts to pinpoint promising practices and extend them to other courts. For example, which practices or procedures have a direct impact on court performance and which do not? How does the performance of specialized juvenile or family courts on these measures compare with the performance of courts of general jurisdiction?

The court performance measures presented below were derived from principles similar to those used for child welfare agencies and focus specifically on child abuse and neglect cases. Although they are the product of a long consensus process and field testing,²⁰ suggestions for improvement are always welcome.

PERFORMANCE MEASURE 1: SAFETY

Goal 1: Children should be safe from abuse and neglect while under court jurisdiction.

Safety measures address the status of children while they are under the jurisdiction of the court. The performance

outcome promoted by these measures follows from the principle of “first do no harm.”

Safety Outcomes Are:

- Children are, first and foremost, protected from abuse and neglect.
- No child should be subject to maltreatment while in placement.
- Children are safely maintained in their homes whenever possible and appropriate.²¹

What Courts Should Measure:

1. Percentage²² of children who do NOT have a subsequent petition of maltreatment filed in court after the initial petition is filed.
2. Percentage of children who are the subject of additional allegations of maltreatment within 12 months after the original petition was closed.²³

PERFORMANCE MEASURE 2: PERMANENCY

Goal 2: Children should have permanency and stability in their living situations.²⁴

Permanency outcomes are closely related to timeliness measures, but also include additional considerations. Assessments of whether the court facilitates permanency include a focus on whether children change placements, whether in the end cases achieve permanent legal status, and whether children reenter foster care due to placement disruption. The permanency measures

presented in this *Guide* and *Toolkit* encourage courts toward the “long view” of the court experience for abused or neglected children. An important challenge for courts addressing the permanency measures is that in order to address them adequately, a court will need to obtain information from partner agencies (e.g., the state child welfare system or private providers who track children placed in foster care).

Permanency is achieved when children are returned to their families without further court supervision, when children are adopted, or when children are placed with individuals who are their permanent guardians.²⁵ Courts are empowered to remove children from home if they are in danger of harm, but also have other alternatives, including removing the alleged perpetrator and placing the child with members of the extended family.²⁶

Permanency Outcomes Are:

- Children have permanency and stability in their living situations.²⁷
- The continuity of family relationships and connections is preserved for children.

What Courts Need to Measure:

1. Percentage of children who reach legal permanency (by reunification, guardianship, adoption, planned permanent living arrangement, or other legal categories that correspond with ASFA) within 6, 12, 18, and 24 months from removal. Specific time lines for this measure should be adapted to jurisdictional timelines.

²¹ Although safety is a concern for both child welfare agencies and courts, the emphasis is different. Child welfare agencies focus attention on reports of abuse or neglect. The court measures discussed here focus on new allegations made while the child is under court jurisdiction. Moreover, courts should be concerned about how often children do return to court with a new allegation after court jurisdiction has been terminated in a previous case.

²² A percentage should not be calculated if the number of cases involved is less than 20. In those instances, the raw frequencies should be reported. Indeed, it is always useful to provide users with the number upon which the percentages were calculated.

²³ The Children’s Bureau of the U.S. Department of Health and Human Services recently changed their definition to recurrence within six months (ACYF-CB-IM-00-11; ACYF-CB-IM-01-01; ACYF-CB-IM-01-07; 45 CFR 1355.34(b)(4) and (5); see also www.acf.hhs.gov/programs/cb).

²⁴ Measures under “permanency” should measure stability as well since federal CFSRs include stability as part of overall permanency. To measure the stability of judicial involvement, the principle at work is consistency of decisions and information as well as the avoidance of loss of relationships.

²⁵ See 42 U.S.C. Sec675(5)(c).

²⁶ *Guidelines for Public Policy, op. cit.*, IV-11.

²⁷ The W.K. Kellogg Foundation’s Families for Kids Program, among other organizations, adds a time dimension to the permanency goal – placement in nurturing, permanent homes within one year. This elapsed time goal will be considered here as an integral part of the measure of permanency because it is a shared goal of courts and social service agencies. The timeliness of court processing, however, will be considered part of Goal 4, discussed later.

²⁸ This measure was originally conceived to cover the scenario during which a child returns home, the court case is closed, and after some time has elapsed, returns to foster care in the custody of the agency. The court may also want to capture information on those cases in which children are returned home under protective supervision, the case remains open, and the child returns to foster care in the custody of the agency after some time has elapsed.

²⁹ American Bar Association, Center on Children and the Law, *State Court Assessments 1995-1998, Dependency Proceedings*, Vol. 2, *Quality of Hearings* (Washington, DC: ABA Center on Children and the Law, 1999), p. 17. See also *Resource Guidelines*, *supra* note 5.

The ABA Center for Children and the Law considers the completeness and depth of child protective hearings to be a major factor in the quality of proceedings.

2. Percentage of children who do not achieve permanency in the foster care system (e.g., court jurisdiction ends because the child reaches the age of majority).
3. Percentage of children who reenter foster care pursuant to court order within 12 and 24 months of being returned to their families.²⁸
4. Percentage of children who return to foster care pursuant to court order within 12 and 24 months of being adopted or placed with an individual or couple who are permanent guardians.
5. Percentage of children who are transferred among one, two, three, or more placements while under court jurisdiction. Where possible, this measure should distinguish placements in and out of a child's own home from multiple placements in a variety of environments.

PROCESS GOALS ARE IMPORTANT TO PERFORMANCE

Process goals focus on how the outcome goals are achieved as well as the quality of services provided. The following goals below focus on two court process goals – due process and timeliness.

PERFORMANCE MEASURE 3: DUE PROCESS

Goal 3: To deal with cases impartially and thoroughly based on evidence brought before the court.

Due process measures address the extent to which individuals coming before the court are being provided basic protections. Due process refers to the right of all parties to participate in court proceedings. Among other things, courts must ensure that family members have notice of the proceedings as well as a fair opportunity to present testimony and express their point of view. These rights apply at all stages of the court process.

The performance goal addressed by these measures is the enhancement of due process by deciding cases impartially and thoroughly, based on evidence brought before the court. This goal encompasses giving each family the individual attention necessary to make effective decisions for the child and assuring that each child receives due process, including effective legal representation. The ideal is that children in similar circumstances should achieve similar results regardless of the jurisdiction in which the case is heard.

The ABA Center for Children and the Law considers the completeness and depth of child protective hearings to be a major factor in the quality of proceedings.²⁹ Quality hearings encompass, in part, notification of parties involved, amount of hearing time allotted, use of court reports, case plans, and findings, and court emphasis on permanency planning. The objective measures of due process proposed below incorporate these concepts of quality proceedings but cannot be complete without

qualitative measures of fairness and equality. Qualitative measures of fairness and equality are discussed in later sections of this chapter.

Due Process Outcomes Are:

- Enhancement of due process by deciding cases impartially and thoroughly, based on evidence brought before the court.

What Courts Need to Measure:

1. Percentage of cases in which both parents receive written service of process within the required time standards or where notice of hearing has been waived by parties.
2. Percentage of cases in which there is documentation that notice is given to parties in advance of the next hearing.³⁰
3. Percentage of cases in which the court reviews case plans within established time guidelines.
4. Percentage of children receiving legal counsel, guardians *ad litem* or CASA volunteers in advance of the preliminary protective hearing or equivalent (Percentage within established time guidelines? Percentage within 0-5 days? 6-10 days? More than 10 days?).
5. Percentage of cases where counsel for parents are appointed in advance of the preliminary protective hearing or equivalent (Percentage within established

time guidelines? Percentage within 0-5 days? 6-10 days? More than 10 days?).

6. Percentage of cases in which legal counsel for children changes (as well as number of changes in counsel if possible).
7. Percentage of cases in which legal counsel for parents changes (as well as number of changes in counsel if possible).
8. Percentage of cases where legal counsel for parents, children, and agencies are present at each hearing.
9. Percentage of children for whom all hearings are heard by one judicial officer (as well as two, three, or more judicial officers if that information is available).³¹

PERFORMANCE MEASURE 4: TIMELINESS

Goal 4: To enhance expedition to permanency by minimizing the time from the filing of the petition or protective custody order to permanency.

Establishing and complying with state and federal guidelines for timely case processing are also important court process performance goals. Limiting the time required to bring litigation to a conclusion limits the exposure of families to emotionally charged issues that can have a detrimental impact on children.³² Long periods of uncertainty and judicial indecision can put pressure on children and families, greatly adding to the strain of foster care. In addition, judicial timeliness is

³⁰ For most courts this may be an "aspirational goal" reflecting best practices.

³¹ To measure the stability of judicial involvement, the principle at work is consistency of decisions and information as well as the avoidance of loss of relationships.

³² Joseph Goldstein, Anna Freud, and Albert Solnit, *Beyond the Best Interests of the Child* (New York: Free Press, 1979). Authors note the importance of considering the child's sense of time.

³³ Two appellate measures are usually included as part of the timeliness goal: (1) Percentage of adjudication, disposition, termination and other judicial decisions that are appealed and percentage overturned on appeal; (2) Percentage of cases in which the results of the appeal are received within 1, 3, 6, and 12 months from the date the appeal was filed. The goals are very important and relevant, but cannot be obtained from trial court case files. For information on how appellate courts can expedite proceedings, see Ann Keith and Carol Flango, *Expediting Dependency Appeals: Strategies to Reduce Delay*, 2nd ed. (Williamsburg, VA: National Center for State Courts, 2002).

It is important not only to capture the total time it takes a child to reach a permanent legal status, but also to capture the time elapsed between events in the court process...

closely related to the goal of permanency. Children can be damaged by “foster care drift” – remaining too long in “temporary” foster homes. Clearly, the length of time required to resolve family issues needs to be limited and reasonable, given the potential harm from delays. Courts need guideposts to help them determine how well they are meeting performance goals.

In some courts, for example, a case can remain in litigation for a year or more after a petition for termination of parental rights is filed, before the trial court makes a final decision. In some courts, it can take up to a year from the date a child is removed from home simply to establish whether or not the child has been abused and neglected and the court has the power to determine who shall have custody of the child. Many courts perform in a far more timely fashion. It is important to capture this dimension of a court’s performance.

It is important not only to capture the total time it takes a child to reach a permanent legal status, but also to capture the time elapsed between events in the court process (e.g., court hearings) so that courts can pinpoint precise sources of delay, and thus improve performance. Courts generally are most familiar with timeliness measures. These measures provide courts with tools to assist them in pinpointing areas where they are doing well and areas where improvement is needed.

Timeliness Outcomes Are:

- Expedition of permanency by minimizing the time from the filing of the petition or protective custody order to permanency.

What Courts Need to Measure:³³

1. Average or median time from filing of the original petition to adjudication.
2. Average or median time from filing of the original petition to disposition.
3. Percentage of cases that are adjudicated within 30, 60, 90 days after the filing of the dependency petition.
4. Percentage of cases that receive a disposition within 10, 30, 60 days after the dependency adjudication.
5. Average or median time from filing of the original petition to permanent placement.
6. Average or median time from filing of the original petition to finalized termination of parental rights.
7. Percentage of cases for which the termination petition is filed within 3, 6, 12, 18 months after the dependency disposition.
8. Percentage of cases that receive a termination order within 30, 90, 120, 180 days after the filing of the termination petition.
9. Percentage of cases for which an adoption petition is filed within 1, 3, 6 months after the termination order.
10. Percentage of cases for which the adoption is finalized within 1, 3, 6, 12 months after the adoption petition.

11. Percentage of hearings (by hearing type) not completed within times frames set forth in statute or court rules. Where possible, the reason(s) for noncompletion should also be captured (e.g., party requesting postponement).

PERFORMANCE MEASURE 5: WELL-BEING

Courts do not have the same extensive role to play in the lives of children and families that child welfare agencies do, and consequently are likely to have fewer outcome goals.³⁴ The court's role in ensuring the well-being of children is more indirect. Although courts do not provide care for children directly, they do have a role in inquiring about the health, medical care, school attendance, and other indicators that children are being properly cared for. These indicators may provide cues of dysfunctional family relationships and cause the family to return to court repeatedly. That being said, it is premature at this time to have courts adopt measures of well-being when consensus does not exist on measures for which courts have direct responsibility, such as safety of children, appropriate removal of children from their homes, successful achievement of permanency, and length of time in foster care. Yet such performance measures are part of a process of continuing improvement, which means that they should be reexamined and refined as their usefulness becomes apparent.

Children's well-being is another dimension of performance measurement that is specified in the Adoption and Safe Families Act (ASFA). In ASFA, children's well-being refers to factors other than safety and permanency that relate to a child's current and future

welfare. Most notably, child well-being under ASFA refers to the child's educational achievement and mental and physical health. Measures of children's educational achievement and mental and physical health are *not* included in this *Guide* and *Toolkit* for several reasons:

- First, neither the federal government nor the social science research community have identified, or achieved consensus on, helpful statistical measures that are specifically related to child welfare cases. By contrast, we were able to adapt measures of safety, permanency, and procedural fairness related to court performance in child welfare cases.
- Second, even if there were clear well-being measures, the judicial branch is not likely to have child well-being statistics readily available. Getting this information requires data exchanges with external entities, which will only become possible *after* the court has developed its own system to measure performance.
- Third, although courts influence children's educational attainment and health only indirectly, they clearly do impact children's safety and permanency.

In the future, it may be helpful for courts to use child well-being measures in analyzing their own performance. To the extent that courts have the responsibility to make sure that the state is providing proper care to children in its custody, it will be useful for courts to know whether those children over whom they have jurisdiction are receiving a good education and are physically and emotionally healthy. If a local court learns, for example,

³⁴ Prevention goals especially may be achieved by child welfare agencies alone without court involvement. For example, Oregon's goal of reducing the number of abused children under age 18 decreased from 12 per 1,000 children to 6 per 1,000 children. Oregon Progress Board, *Oregon Benchmarks: 1993 Report to the Legislature* (Salem: Oregon Progress Board, 1993).

**In the future,
it may be
helpful for
courts to use
child well-being
measures in
analyzing their
own
performance.**

³⁵ The data elements required to produce the core measures listed above are outlined in the *Packard Court Performance and Judicial Workload Toolkit* and are available from the partner organizations upon request.

that children in court-supervised foster care are substantially behind educationally, the court may decide to ask more penetrating questions about children's educational attainment. The court may decide to demand more documentation concerning the child's education, may instruct guardian *ad litem*s to check into children's educational progress, and may even decide to join in meetings with school officials to discuss the educational needs of children in foster care and how best to address them.

Accordingly, once useful well-being measures have been developed for child welfare cases, at least some courts will want to include them in their own system for performance measurement. Data to support these measures, however, will primarily have to come from sources external to the court.

CAN COURTS AND AGENCIES ACTUALLY PROVIDE THESE MEASURES?³⁵

It is one thing to construct measures to demonstrate improvement in court performance for child abuse and neglect cases. It is quite another task to field test the instruments to determine the extent to which sites can provide these measures given their current operating systems. It is quite a confirmation of the usefulness of the proposed measures if states and agencies are already collecting the data elements required to construct the measures. To the extent that there is commonality among the data elements and measures already being reported, it shows that sites themselves have recognized the importance of knowing how well their courts are performing in these areas. Sites contemplating

establishing or revising court performance measures should receive encouragement from the fact that other sites can produce the performance measures, without major modifications to their existing information systems.

Generally, the conclusions we reached after conducting six project site visits about sites' abilities to provide court performance measures are:

1. It is more difficult to obtain measures that require both court and child welfare agencies to exchange information.
 - Measures of child safety, the most critical indicator for both courts and child welfare agencies, require close collaboration between these respective information systems.
 - Sharing of information is essential for the analysis of permanency measures, especially to determine why children reenter foster care after being reunified with their families or permanently placed.
 - Changes in placement require courts to receive information from child welfare agencies.
2. Due process measures are more process than outcome oriented and should be supplemented by qualitative measures that go beyond assignment of counsel and service of process, but attempt to discern the quality of the hearings and the treatment of the parties.

- Due process measures were not readily available in electronic form from any of the three project test sites.
- Courts were most likely to be able to construct measures of timeliness, perhaps because information about time between events is necessary for caseflow management.

DECIDING WHICH PERFORMANCE MEASURES TO USE

Despite the time and effort that went into constructing the recommended court performance measures presented in this *Guide* and *Toolkit*, a site should critically review them before accepting the measures. The performance measures will work best if each site examines each measure individually and determines:

1. That the measure is indeed important, needed, and useful;
2. That the information to produce the measure exists or that proxy measures are possible;
3. That issues such as privacy of records can be overcome; and
4. That the usefulness of the information justifies the cost of obtaining the measure.

METHODS FOR OBTAINING COURT PERFORMANCE MEASURES – THE IMPORTANCE OF TRIANGULATION

Typically, any single data collection method will not be completely satisfactory for a comprehensive performance and workload evaluation. When at all possible, it is always better to use several different collection methods and sources of data. The process for court performance and judicial workload assessment outlined in this *Guide* and *Toolkit* strongly recommends that jurisdictions use multiple methods and data sources to inform their self-assessments. This approach, known as “triangulation,” allows you to combine strengths and correct some of the deficiencies of any one source of data.

Triangulation is a term used to denote the process of building checks and balances into a research design through multiple data collection strategies and is aimed at increasing rigor. For example, **Data Triangulation** involves the use of a variety of data sources to study the same phenomenon (e.g., interviewing people in different positions about court performance issues).

Methodological Triangulation involves the use of multiple methods to study a single issue (e.g., using interviews, questionnaires, focus groups, court observation, and analysis of management information systems to study court performance and judicial workload).

It is often desirable to measure or assess the same result based on a number of different data sources, as well as through a number of different evaluation designs. This is

the case because any given evaluation method may be open to threats to internal validity, where alternative explanations cannot be entirely ruled out or accounted for. Consequently, complementary strategies can be an effective means of ruling out rival explanations for observed outcomes. For this reason, it is desirable to address assessment or evaluation issues from a number of different perspectives, using multiple lines of evidence to lend greater credibility to the evaluation findings. When independent strategies relying on different data sources and different analytical techniques converge on the same conclusion, you can be reasonably confident that the conclusions are reliable. Of course, if individual strategies lead to varying conclusions, the situation is somewhat problematic. Nevertheless, this result is preferable to carrying out a single strategy and unknowingly drawing conclusions that would be contradicted by a different strategy.

For example, consider an assessment of the impact of a mediation program on the child protection hearing process. This issue could be addressed in a number of different ways. One strategy would be to survey system professionals about the success of the mediation program. This strategy would provide useful information about stakeholders' perceptions of the operation and overall value of the mediation program. However, for a number of different reasons, including a possible desire to see continued funding of the program, survey respondents might exaggerate the effect or impact of the program on case processing. This problem would indicate the need to investigate the effects of the program in other ways. For instance, survey responses could be supplemented with a review of mediation case records, recording time frames from petition filing, mediation, and case closure. These rates could then be compared to case-processing time frames for a sample of similar cases that did not receive mediation.

Attention to multiple methods and data sources involves a mix of both **quantitative and qualitative** approaches to measurement. Quantitative approaches to measurement involve heavy use of numerical data and analysis methods (e.g., data captured in an automated management information system). Qualitative approaches, on the other hand, focus on providing description and interpretation. The use of qualitative methods can provide you with rich, firsthand information on questions such as how procedures and programs are implemented, the patterns of interactions between stakeholders, the kind of day-to-day problems that are confronted by system stakeholders, and the obstacles faced to achievement of best practice and court improvement goals. Qualitative data can provide stakeholder suggestions for future court improvement efforts. Qualitative inquiry emphasizes “thick description” (i.e., obtaining data that illuminate everyday patterns of actions), and can be a critical tool for gaining “buy-in” for evaluation findings – individuals interviewed or surveyed, or who have participated in focus groups, will feel that their voices are reflected in the study's outcomes.

QUANTITATIVE AND QUALITATIVE COURT PERFORMANCE MEASUREMENT METHODS – A BRIEF OUTLINE OF STRENGTHS AND LIMITATIONS

Methods used to collect data must be selected on the basis of the nature of the data required and the sources available. Each of the methods suggested in this *Guide* and *Toolkit* for obtaining court performance data are introduced briefly below, with a discussion of their associated strengths and limitations. The instruments

associated with each method are presented in detail in the *Packard Court Performance and Judicial Workload Assessment Toolkit*, including instructions for using the instruments.

For the ease of presentation and the sake of discussion, each data collection method is presented separately. However, in the context of a court performance and judicial workload assessment, quantitative and qualitative methods should be used *together* to support the various evaluation research strategies employed.

CASE FILE REVIEW

A file review is a data collection method aimed at discovering preexisting, objective data that can be used in the evaluation. The data collected in a case file review can be quantitative (e.g., numerical data related to time frames for court events) or more qualitative (e.g., presence of parties at specific hearings or completeness of orders). File data may be retained by a computerized management information system or exist only in a hard-copy file. File reviews can be useful for:

- Providing background data and information on the hearing process, timeliness, and the court environment, thus putting court performance and workload results in context.
- Providing measures of court performance when such information is not available from computerized management information systems.
- Producing a useful framework and basis for further data gathering.

- A file review may establish the population (sampling frame) from which the survey sample is to be drawn. Background information from the files (e.g., demographic characteristics, petition allegations) may be used in designing the most powerful sample, and in preparing the interviewer for an interview. Asking for information on a survey that is already available in files is a sure way of discouraging cooperation.

There are, however, certain problems that may occur with a file review:

- Files may be incomplete or unavailable.
- File reviews are time and labor intensive, making them less economically feasible.

DIRECT OBSERVATIONS

“Seeing is believing,” as the old saying goes; direct observation generally provides more powerful evidence than that which can be obtained from other sources. The results of observations, recorded through some standardized technique (e.g., using a court observation checklist or form), may have a powerful impact on the reader of an evaluation report. Observation involves selecting, watching, and recording objects or activities that play a significant part in the event being evaluated. The observed conditions can then be compared with some preestablished criteria and the deviations from these criteria analyzed for significance. Observational data describe the setting of an event, the activities that take place in the setting, the individuals who participate in the activities, and the meaning of those activities to

“Seeing is believing,” as the old saying goes; direct observation generally provides more powerful evidence than that which can be obtained from other sources.

the individuals. With respect to court hearing observation, the technique involves on-site visits to observe a representative sample of hearings firsthand and to record what takes place in a standardized way.

Observations should be written up immediately after the visit and should include enough descriptive detail to allow the reader to understand *what* has occurred and *how* it occurred. Descriptions should be factual, accurate, and thorough, without being filled with irrelevant items. Observational data are valuable in evaluations because evaluators and users can understand activities and effects through detailed descriptive information about what occurred.

Strengths and weaknesses of observational techniques:

- Permits a better understanding of the object of study – creates a more complete picture of court functioning.
- Allows you to move beyond the selective perceptions gained through such means as interviews. Evaluators, as field observers, will also have selective perceptions, but by making their own perceptions part of the data available, they may be able to present a more comprehensive view of the activity being studied.
- Provides evaluators with the chance to see things that may escape stakeholders' notice, or issues that they are reluctant to raise in an interview.

Court hearings may involve routines that individuals take for granted; subtleties may be apparent only to those not fully immersed in these routines. This often

makes it possible for an outsider, in this case the evaluator, to provide a “fresh” view. Similarly, outsiders may observe things that participants and staff are unwilling to discuss in an interview. Thus, direct experience with and observations of hearings will allow evaluators to gain information that might otherwise be unavailable.

- The reliability and validity of observations depend largely on the skills of the observer and on the observer's awareness of any bias he or she brings to the task.

Another person carrying out a similar set of on-site observations may observe the same hearing differently. This implies limits to both the internal and external validity of the direct observation data. However, if observers are well-trained and use a pretested, standardized observation coding tool, such as the one included in the *Packard Toolkit*, these threats to validity can be minimized.

- Individuals may behave quite differently from their usual behavior patterns, if they know that they are being observed by an evaluator.

Evaluators must be sensitive to the fact that individuals may act differently if they know they are being observed. Appropriate steps to prevent this problem from occurring, or to account for this effect, should be taken.

- Comprehensive court observations (i.e., observations that sample sufficient numbers of judges and hearings)

can be labor and time intensive. (They cannot be done in one visit and may require someone on-site to observe rarely occurring events.)

SURVEY METHODS

Surveys are systematic ways of collecting primary data – quantitative, qualitative, or both – from persons associated with the event or process under study. The term “survey” refers to a planned effort to collect needed data from a sample (or a complete census) of the relevant population. The relevant population is composed of those persons from whom the data and information are required. When properly conducted, a survey offers an efficient and accurate means of ascertaining the characteristics of almost any population or activity of interest. Surveys can take the form of interviews or questionnaires.

The first and most fundamental step in preparing a survey is to identify, as precisely as possible, what specific information will address a given evaluation issue. Next, the evaluator must consider economy and efficiency. There is always a temptation to gather “nice-to-know” information. The evaluator should realize that defining the scope and nature of a survey determines in large part its cost and that collecting “extra” data will add to the total cost.

Surveys that have not been properly pretested often turn out to have serious flaws when used in the field. Pretests should involve testing both the questionnaire and procedures to be used in conducting the survey.

Pretesting will provide information on the following:

- Clarity of questions – Is the wording clear? Does every respondent interpret the question in the same way? Does the sequence of questions make sense?
- Response rate – Is there any question that respondents find objectionable? Does the interview technique annoy respondents? Do respondents refuse to answer parts of the questionnaire?
- Time and length – How long does the questionnaire take to complete?
- Survey method – If the survey is conducted by mail, does it yield an adequate response rate? Does a different method yield the required response rate?

Strengths and weaknesses of survey methods:

- A survey is a very versatile method for collecting data from a population.
- Using a survey, one can obtain attitudinal and perceptual data on almost any aspect of a program and its results. The target population can be large or small, and the survey can involve a time series of measurements or measurements across various populations.
- When properly conducted, a survey produces reliable and valid information.
- Surveys require expertise in their design, implementation, and interpretation. They are easily

The first and most fundamental step in preparing a survey is to identify, as precisely as possible, what specific information will address a given evaluation issue.

misused, resulting in invalid data and information. Survey procedures are susceptible to a number of pitfalls that threaten the reliability and validity of the data collected: sampling bias; nonresponse bias; sensitivity of respondents to the questionnaire; interviewer bias; and coding errors. Each potential problem must be controlled for and the survey process must be rigorously controlled for quality.

EXPERT OPINION FOCUS GROUPS OR DELPHI GROUPS

Expert opinion, as a data-gathering technique, uses the perceptions and knowledge of experts in given functional areas as evaluation information. Essentially, this method consists of asking a group of experts in a given field for their opinions on specific evaluation issues. Evaluators use this information to determine outcomes. Eliciting opinions from experts is really a specific type of survey, and all the comments described under the survey section are relevant here. Note, however, that expert opinion is a method best suited to supplementing (or replacing) data in the absence of more objective indicators. It does not refer to the use of experts on the evaluation or assessment team, but rather to the use of experts as a source of data for addressing evaluation issues. Expert opinions can be collected and summarized systematically, though the results of this process will remain subjective.

Strengths and weaknesses of focus groups or Delphi groups:

- Expert opinions can be used to carry out measurements in areas where objective data are deficient. It is a relatively inexpensive and quick data collection technique.
- There may be a problem in identifying a large enough group of qualified experts if the evaluator wishes to ensure statistical confidence in the results.
- Like any subjective assessment, expert opinion presents a credibility problem.
- Disputes over who the expert participants were and how they were chosen can easily undermine the best collection of expert opinion.

SUMMARY OF METHODS

In the context of a court performance study, each of the quantitative and qualitative methods should be considered as measurement tools. Observations and file reviews help define the context of the activities under study and suggest plausible ways of attributing observed results to a given process, procedure, or program. Attitudinal data obtained from surveys gives you the aggregate opinion of the target population about actual results. Such data also builds “buy-in” for the findings of the study.

While you should be aware of the potential subjectivity of data obtained through more qualitative methods, such

as interviews, observations, and expert opinion focus groups, it is not necessarily a disadvantage of these methods. Such qualitative data collection methods are the best ways to generate holistic, in-depth information. Used with quantitative data, qualitative data are quite effective in verifying the link between a program, procedure, or process and its observed results. Typically, any single data collection method will not be completely satisfactory for a comprehensive performance and workload evaluation. When at all possible, it is always better to use several different collection methods and sources of data.

CONCLUSION

The question for outcome-oriented measures is always the extent to which they measure whether the results desired are being achieved. Outcome measures change the orientation from “What efforts are being made?” and “What work is being performed?” to “What results are being achieved?” Outcomes are the ultimate measure of success because they force an objective appraisal – did the intervention make a difference?

Because there is a resistance to measurement, some of it justified, performance measures themselves, to be credible, should be evaluated against the following standards:

- **Appropriateness and Validity** – Do measures adequately reflect what is important to measure and the extent to which information needs are being met?
- **Completeness** – Do the set of measures taken together cover all of the objectives?

- **Comprehensibility** – Are the measures understandable?
- **Timeliness** – Are performance results available soon enough for managers to respond, rather than being an “autopsy” showing what went wrong?
- **Cost-Effectiveness** – Can measures use data already being collected or are costs of collection reasonable?
- **Visibility** – Can the measures be tracked by those being measured?
- **Confidentiality** – Are there concerns for client or user privacy?
- **Comparability** – Are measures comparable within states and across states? Can performance be measured across court and child welfare systems?³⁶

The measures presented in this *Guide* are intended to be developmental – refined as more information is known about the attributes associated with a model process for the handling of child abuse and neglect cases. As courts and child welfare agencies begin to standardize the measures used to evaluate their own performance, software developers will have the incentives necessary to provide these measures as a by-product of operational software.

Moreover, outcome measures need to be part of a process of continuing improvement, which means the continuing modification and rejection of individual measures that do not measure what they purport to measure or do so inefficiently.³⁷ Performance measures will be more readily

³⁶ See Harry Hatry et al., *Program Analysis for State and Local Governments*, 2nd ed. (Washington, DC: Urban Institute, 1987), reprinted in the National Association for Court Management's *Holding Courts Accountable: Counting What Counts* (Williamsburg, VA: National Association for Court Management, 1999).

³⁷ See E. Finkle, *Performance Measurement Minefield*, *Public Innovator* 43 (1995): 1-2.

Outcomes are the ultimate measure of success because they force an objective appraisal – did the intervention make a difference?

accepted if they are used positively, that is, to reward improvement rather than to retroactively punish failure. They are more likely to be accepted if the focus is on helping states improve services to maltreated children, *not* to evaluate the performance of individual judges or child welfare administrators. Measures that provide timely feedback to courts and child welfare agencies will encourage cooperation to make further improvements in the benchmarks.

Chapter 3

Assessing Judicial Workload

Workload measures are intended to capture the resources necessary for courts to achieve their performance objectives. They enable states or counties not only to distribute the current number of judges more equitably, but also to determine how many judges each court needs to achieve a high level of performance. Establishing appropriate workloads objectively will strengthen the courts' requests for resources needed to process child abuse and neglect cases. Objective methods and measures for appropriate judicial workloads enable courts to approach their legislatures with well-supported funding requests for child welfare cases and to demonstrate objectively the consequences of a lack of resources on specific areas of performance. Objective workload assessment also will help child advocates determine whether courts are able to meet the needs of children and families, thus helping to ensure the safety, permanency, and well-being of children and, at the same time, fulfilling the letter and spirit of the federal Adoption and Safe Families Act (ASFA).

Establishing appropriate workloads objectively will strengthen the courts' requests for resources needed to process child abuse and neglect cases.

WHAT IS BEING MEASURED?

The amount of time judicial officers currently devote to all aspects of dependency cases (including non-case-related activities) and how much judicial time, and how many judges, the court needs to perform in compliance with national standards. Judicial workload needs in abuse and neglect cases are affected by the specific and individualized character of abuse and neglect cases. Along with the rapidly developing and changing nature of abuse and neglect proceedings, courts have fallen behind in adjusting to additional and new workload requirements. Child welfare litigation requires an updated approach to workload assessment and should take the following into account:

- The new types of hearings now required by federal and state law;
- The increased frequency of certain types of these new hearings, whether imposed by law or recommended by national standards;
- The growing number of issues that state laws and national standards now require judges to address in each type of hearing (e.g., agency efforts to locate missing parties, involvement of fathers, reasonable efforts to prevent placement, child support, timely assessments and services, and many others);
- The increasing number of attorneys, parties, and participants in abuse and neglect cases (e.g., putative fathers, noncustodial parents, relatives, foster parents, attorneys for noncustodial and putative parents,

substance abuse treatment providers, and others);

- The increasing procedural protections for parties (e.g., mandatory judicial findings and notice); and
- The additional stricter deadlines for hearings, decisions, and petitions.

BEFORE THE WORKLOAD ASSESSMENT BEGINS

1. IMPACT OF RESOURCES ON COURT PERFORMANCE

One reaction from courts not faring well on measures of court performance is predictable - they lack the resources necessary to perform well. There are not enough judges, quasi-judicial officers, or court support staff to do the quality job that child abuse and neglect cases deserve. Consequently, hearings may be held without proper preparation, too little time may be devoted to each hearing, and the quality of decision making may be degraded in other ways.

Conversely, some courts appear to perform well given roughly the same resources as the courts not performing as well, leading some to conclude that resource levels do not affect performance. Our conclusion is that courts cannot perform well without a minimum set of human and capital resources, but that resources alone do not guarantee good performance. Smaller workload alone will not ensure good performance, although large workloads will surely prevent it.

Workload is only one determinant of court performance. “No matter how great the stature or lavish the resources, to be effective, courts must be well-managed.”³⁸ Workload needs to be linked with other types of management reforms to yield better outcomes for children. Consider the following influences:

- **Case Management:** Effective case management reduces unnecessary court delays and continuances and reduces needed court time for each case by setting firm hearing dates, preparing for hearings in advance, and avoiding duplicative hearings and hearing preparation. A court that permits double booking of hearings, is casual about allowing continuances, and does not track adherence to deadlines will not benefit from a reduction in workload the way it could without implementing more effective case management.
- **Legal Culture:** Legal culture can affect workload. Where the culture is overly adversarial, with attorneys contesting issues to a degree that is ultimately unproductive to their clients, workloads tend to increase. Legal culture not only can affect the number and length of contested hearings, but also the length of routine uncontested hearings.
- **Courtroom Management:** If parties are efficiently ushered in and out of the courtroom, the judge has more time available for the actual hearings. Courtroom management may require, among other things, more, or better trained or more efficiently deployed, court staff.

³⁸ U.S. Commission on Child and Family Welfare, *Parenting Our Children: In the Best Interest of the Nation* (Washington, DC: U.S. Commission on Child and Family Welfare, 1996), Mary R. Cathcart, Chair.

“No matter how great the stature or lavish the resources, to be effective, courts must be well-managed.”

States need workload assessments to achieve the best use of scarce resources – judges, judicial officers, and court support staff.

- **Judicial Expertise:** Experience and specialization affects case-processing time. With individual calendars and judicial specialization, judges require less time to prepare for each hearing, and cases can be concluded earlier.
- **Judicial Support Staff:** Judicial productivity, and hence the need for new judges, depends substantially on the effectiveness of trial court support staff. Without the proper type and level of support, as well as efficient deployment of staff, judges may be performing judicial tasks less efficiently or may be performing tasks that could be delegated to others. For example, judicial support staff can save judges' time by preparing court orders, setting cases, responding to telephone calls, responding to routine messages, and conducting legal research.
- **Pre-Hearing Meetings:** These meetings, in which the judge is not present, can help the parties more sharply focus on issues and speed the resolution process. They can help parents and foster parents have a better understanding of the issues before the court, thus reducing pressure on the judge to explain the proceedings. These meetings can reduce the number and duration of contested proceedings and avoid disagreements in court.
- **Mediation:** Use of mediation can reduce the number and duration of contested proceedings and also the time needed for the judge to explain the proceedings in court.

- **Family Group Conferences:** Family group conferences can avoid the need for court proceedings in some cases, reducing the demands on the court. If there are family group conferences *after* court proceedings have begun, they can have a similar impact on judicial workload as mediation.

2. CURRENT METHODS OF EVALUATING RESOURCE NEEDS

Alternatives to determining the level of sufficient resources range from subjective perceptions to sophisticated weighted caseload techniques. Having objective measures of the number of judges and court support staff needed to properly handle the child abuse and neglect workload may assist decision makers in obtaining the resources necessary to protect children. Certainly the converse is true. Although empirical information alone will not guarantee a favorable staffing decision, it is unlikely in these times of scarce resources that courts will obtain the judges and staff necessary to process dependency cases *without* a strong, empirically based demonstration of need.

States need workload assessments to decide how to achieve the best use of scarce resources – judges, judicial officers, and court support staff. A workload assessment is usually triggered by the belief that the current level of resources is inadequate and that existing resources and technology are not being deployed optimally. Workload assessment is a structured process that allows judges, judicial officers, and court managers to assess the reasonableness of current case-processing practices. Over time, it is often the case that workload rises more quickly

than judicial resources so that the judiciary is increasingly being asked to do more with less. As explained above, this is especially true in abuse and neglect cases because of the rapidly expanding judicial responsibilities in these cases both on and off the bench. As a result, the average amount of time judges currently have to spend in child abuse and neglect cases usually are not sufficient to provide fair and equitable service to the public.

Workload assessments help courts determine what changes in staffing and procedures they can make with the least detrimental impact on children and families and estimate the resources necessary to make major gains in court performance. Moreover, workload assessment provides objective and standardized assessments of resource needs among jurisdictions that vary in population and caseloads.

3. AUDIENCE FOR THE ASSESSMENT

No method of workload assessment will be successful unless it is credible not only to the courts, but also to the bodies that fund them. In some states with a unified financial structure, all judges and court support staff are funded by the state. In those states, the state legislature is the funding body that must be persuaded to add new judgeships or court support staff. In those states, the competition for funding extends to *all courts* and covers *all types of cases*.

In some states, county-funding bodies determine court staffing and must choose among the courts within a particular jurisdiction. Among the questions they may have to decide are: Should a judgeship be assigned to a

criminal court, a drug court, or a dependency court? What criteria can be used to decide? There may be competition among a county's courts regarding how much judicial time will be set aside for particular courts and particular different types of cases. Under these circumstances, a workload assessment can help identify the courts and case types with the greatest need.

In the remaining states, a combination of funding sources is used, with the state often paying the salaries of judges, and the county responsible for the salaries of court support staff and for facilities. The workload assessment procedure described in this chapter below can help regardless of how the judiciary is funded and can be modified to assess the judicial resources for dependency cases only.

THE WORKLOAD ASSESSMENT: STEP-BY-STEP ³⁹

Weighted caseload is a method used to translate court cases into workload. While case counts have a role in determining the demands placed on state courts, raw, unadjusted case-filing numbers offer only minimal guidance as to the amount of judicial work generated by those case filings. Moreover, the inability to differentiate the work associated with each case type can lead to the misperception that equal numbers of cases filed for two different case types result in equivalent workloads. Cases vary in complexity, and different types of cases require different amounts of time and attention from judges, judicial officers, and court support staff. For example, a recent workload study conducted in California found that a typical dependency case required 224 minutes of judge time from filing to resolution, including post-judgment

³⁹ For a general discussion of workload assessments, see Victor E. Flango and Brian J. Ostrom, *Assessing the Need for Judges and Court Support Staff* (Williamsburg, VA: National Center for State Courts, 1996).

⁴⁰ National Center for State Courts, "California Judicial Workload Assessment: Final Report" (August, 2001).

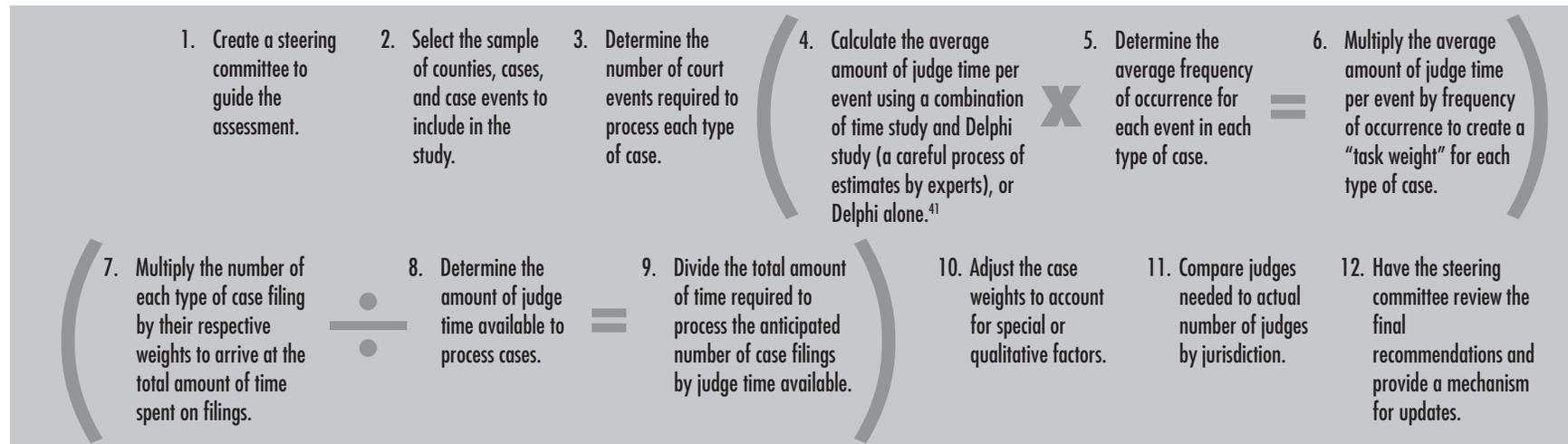
⁴¹ A pure time study (only measuring current time for hearings and not taking into account the time needed for *best practice*) is often appropriate for many types of litigation, especially those that have not drastically changed in recent years. On the other hand, pure time studies are seldom appropriate for analysis of workload in child abuse and neglect litigation. Due to rapid increases in courts' responsibilities for child abuse and neglect cases (more types of hearings, more issues to address, additional parties and participants in each case), few courts currently have workloads that allow them to comply fully with child abuse and neglect laws and to follow best practices in child abuse and neglect cases.

activity.⁴⁰ In contrast the typical delinquency case required 60 minutes. Consequently, a judge could complete nearly four delinquency cases in the time he or she typically took to complete one dependency case. Workload measures determine how many judges and judicial officers are needed to resolve all cases coming before the court and assess whether those resources are being allocated and used prudently.

The workload assessment is a process of converting caseloads into workload using time as a proxy for workload. It is based on the assumption that the more time it takes to process a case, the more work is involved. A workload assessment answers the question, "What is the typical amount of time that a judge or court support staff actually spends on a type of case from filing to final resolution?"

The following diagram is intended to be a *guide* to conducting a workload assessment. Anyone wanting to do an assessment should consider each of the steps even if resources do not permit conducting a full-blown assessment. The heart of the assessment is determining the amount of time each case event takes, and should take, and then calculating the amount of time each case takes to resolve. Multiplying the average time per case by the number of cases yields the total amount of time necessary to dispose of all of the cases, and dividing that time by the amount of judge time available links the workload to the number of judges. In brief, how does the workload compare to the judge's available time to process cases?

Conducting a Workload Assessment



There are multiple ways to determine how much time a case event takes and should take, ranging from making measured observations in courts, recording time on a log, and estimating time based upon expert opinion. The procedure recommended below is a best practice based upon experience, and recommends a combination of time logs and judicial estimates. A site may choose to do one or the other alone based upon resource constraints, but should at least consider the steps recommended here.

STEP 1: CREATE A STEERING COMMITTEE TO GUIDE THE ASSESSMENT

Before considering the procedure to calculate the time it takes by either judges or court support staff to bring a case from filing to final resolution, it is important that a steering committee of decision makers be selected to make decisions about policy questions affecting the study. Policy issues include the length of the study, the types of cases included, the key events in each case, and how to collect both case-related and non-case-related time data. To achieve appropriate case weights for child abuse and neglect cases, it is important for there to be members of the steering committee who are very knowledgeable about abuse and neglect litigation. Because of the unique nature of child abuse and neglect cases, which demand from judges both on-the-bench case-related activities and off-the-bench activities in order to comply with federal and state guidelines for case processing and best practice implementation, the committee should include members who have substantial experience and a deep understanding of these cases.

STEP 2: SELECT THE SAMPLE OF CASES AND COURTS

The confidence in conclusions drawn from any research depends on the adequacy of the sample taken. If chosen properly, a sample will closely approximate the information derived from a study of all cases from all counties.

SAMPLING SITES

Sample selection involves selecting a set of courts, court districts, or circuits for analysis and then determining the variety of case types and case events. Obviously, the most straightforward strategy would be to draw a random sample from all courts, but this is expensive and unnecessary. It is possible to get the same results as a statewide sample, at a much lower cost, by choosing a representative sample that is stratified to ensure that large, medium, and small counties are represented.

Other criteria for site selection include the following:

GEOGRAPHIC DIVERSITY

The types of cases filed in court are likely to differ in rural and urban areas. A weighted caseload study should determine whether urban counties have a different case mix than smaller, more rural counties and whether the differences are sufficient to justify separate weights. Moreover, studies have shown that judges in urban courts spend their time differently than judges in more rural areas.

Multi-judge courts may require judges to spend more time on administrative and coordination activities, whereas single-judge courts in rural areas may necessitate judges to spend more time traveling.

In states with large rural populations, access to justice might require service to a dispersed population, even if the caseload volume would not ordinarily justify a court. These low-volume courts are necessary to provide the public access to court without having to travel extremely long distances. Consequently, public access to the courts, not caseload, is the primary criteria for workload assessment. In those states, it may be possible to include only the single-judge jurisdictions with the largest caseloads in the pool from which the sample is drawn. These are the jurisdictions that will most likely require a second judge. Reducing the number of jurisdictions in the sample will reduce the cost and burden of conducting a weighted caseload study.

SIZE OF COURT

Case-processing procedures may vary between large and small courts. Size is operationally defined here as number of judges. Multi-judge courts may require judges to spend more time on administrative and coordination activities, whereas single-judge courts in rural areas may necessitate judges to spend more time traveling. Stratification by size of court ensures that all courts are represented in the sample, and thus permits the results of this research to be generalized to the entire state.

CASE-PROCESSING TIME

The average time required to process cases will be greater in inefficient courts. Consequently, weights derived from these courts will overestimate the need for judges. Particularly well-run courts are likely to be more efficient and effective than others and should be strongly considered for inclusion in the weighted caseload study.

Deciding which courts efficiently process child abuse and neglect cases can be more complicated compared to other types of litigation. For example, processing times for criminal cases might be measured simply from the filing of the criminal complaint to the completion of sentencing. By contrast, as explained below, it is not possible to measure the efficiency of case processing in child abuse and neglect cases by merely considering the average time between child abuse and neglect petitions and the disposition of such petitions.

While the average time from an abuse or neglect petition to disposition is one indicator of efficient case processing, there are other equally or more important indicators. In short, an efficient court should have:

- A relatively short average time from the original abuse or neglect petition to disposition.
- A relatively short average time from disposition to the filing of a termination of parental rights (TPR) petition.
- A relatively short average time from the filing of a TPR petition to the trial court TPR decision.
- A relatively short average time from a child's removal from home (and notice of such removal to the court) to the time of case closure following the child's return home, adoption, or legal guardianship.
- A reasonable proportion of cases in which children in foster care are reunified or adopted.

Unfortunately, all of these statistics are not readily available in many states. Nevertheless, to select courts that efficiently process child abuse and neglect cases it is important to at least make educated estimates. County statistics from the state child welfare agency may help – to identify jurisdictions where the average time to adoption of children in foster care is relatively short, for example. Further, it is helpful to select courts that most closely follow nationally accepted best practices, such as early notice to all parties, early appointment of counsel, case-specific reasonable efforts findings, and other best practice features described in the *Resource Guidelines*.⁴²

QUASI-JUDICIAL OFFICERS

Courts that make efficient use of quasi-judicial officers (whether they are called commissioners, magistrates, masters, or referees) should require fewer judges than courts without such support. For that reason, use of quasi-judicial officers who complete any task that would otherwise be done by judges should be included as a variable in any weighted caseload study.

AUTOMATION

Courts that use automated information systems may be gathering the type of information that enhances management and operational control of the court. Automation holds the promise as a ready source for the detailed caseload information required by the weighted caseload technique. A related point is that automation encourages data to be collected in a consistent and comparable manner throughout the state, thus making court-by-court comparisons possible and equitable.

SAMPLING CASE TYPES

The question of how many case types to include in a weighted caseload study is one of balance. The more case types used in the study, the more detailed and precise a weighting scheme will be. The more case types included, however, the larger the sample size required, and hence the larger the burden and cost of the weighted caseload study.⁴³

State-level and large-county policy makers are often interested in studies that include the major case types so they can determine where the needs for judges and court support staff are the greatest. Thus, some weighted caseload studies may not address abuse and neglect litigation. Similarly, if a state has specialized juvenile or family courts, the study might encompass only family cases to answer the question of how to allocate those resources. It is even possible to sample only one type of case, say child abuse and neglect, in one jurisdiction if that is the information required to persuade funding bodies.⁴⁴ Where policy makers suspect that judicial resources may be misallocated in a particular type of case, they may decide to sample only that type of case, say child abuse and neglect.

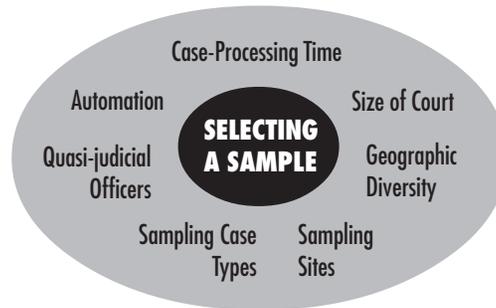
STEP 3: IDENTIFY THE NUMBER OF CASE-PROCESSING EVENTS REQUIRED TO PROCESS EACH TYPE OF CASE

The units of analysis in a weighted caseload study are the case events – the set of activities that make up a case. All potential judicial activities associated with a particular type of case are classified into a set of “event” categories.

⁴² *Supra*, note 5.

⁴³ In addition, when more case types are included in the workload study (for example, other family matters, criminal cases, etc.), the workload needs specific to child abuse and neglect cases may be oversimplified. In a study of many different case types, it becomes more difficult to capture accurately the times needed for each of the unique hearing types in child abuse and neglect cases and more difficult to get accurate estimates of the time needed to fulfill best practice for each hearing type. This problem can be especially serious, given the recent expansion of judicial responsibilities in child abuse and neglect cases.

⁴⁴ This approach was first reported in David Steelman, H. Ted Rubin, and Jeffery Arnold, *Circuit Court of Cook County, Illinois Juvenile Division Judge Workloads and Judgeship Needs* (Chicago: Circuit Court of Cook County, 1993) in a study funded by the State Justice Institute and Cook County. See also: David Steelman et al., *Preliminary Assessment of Judicial Workloads for Juvenile Dependency Cases in Santa Clara County, California* (Denver: National Center for State Courts, Court Services Division, 2000). Reproduced in *Symposium on Improving Outcomes for Abused and Neglected Children* (Williamsburg, VA: National Center for State Courts, 2000) as part of a grant from the David and Lucile Packard Foundation to the ABA Center on Children and the Law, the National Center for State Courts, and the National Council of Juvenile and Family Court Judges.



Common events include:

- Preliminary proceedings
- Hearings and motions
- Trial
- Disposition
- Post-judgment activity
- Case-related administration

Note that these are all *potential events*, because all cases do not have all events. Once the individual events are identified, a sample of each event is examined to determine the average amount of time required to accomplish each one. If the analysis is only being done on child abuse and neglect cases, the analysis can be more detailed. We recommend constructing a flowchart showing all of the case events and how they occur chronologically. In abuse and neglect cases, some of the most typical events are:

- Shelter care hearing (also may be known as emergency removal hearing, temporary protective hearing, detention hearing, emergency hearing)
- Pre-adjudication motions and hearings

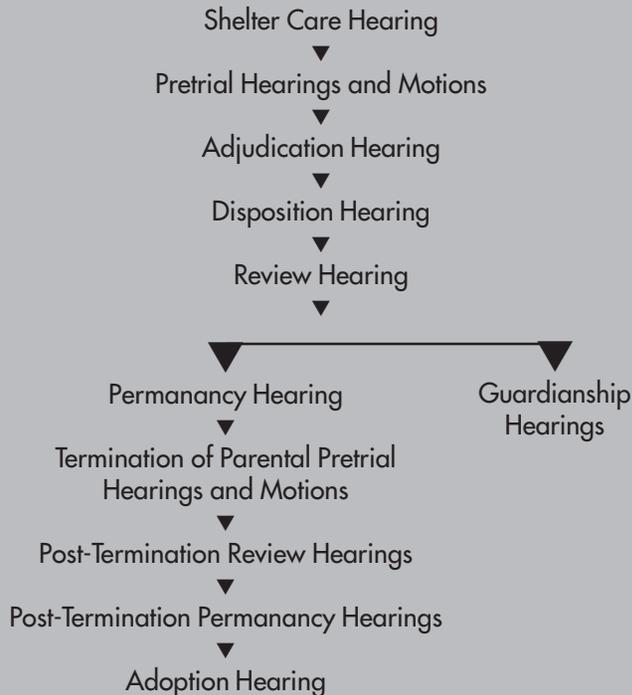
- Adjudication (also may be known as trial or fact-finding hearing)
- Disposition hearing (may be held immediately following the adjudication hearing as a bifurcated adjudication-disposition hearing)
- Review hearings
- Permanency hearings
- Pre-termination of parental rights motions and hearings
- Termination of parental rights hearing
- Post-termination review hearings
- Adoption hearing

Research should be done to establish required time frames and policies on continuances. The analyst may also want to compare briefly how the local process compares with the process envisioned in the Adoption and Safe Families Act of 1997 (ASFA).

How many case events should be sampled? If processing time for certain case events do not vary much, the sample can be relatively small and still provide a good estimate. Conversely, if processing times vary dramatically, a larger sample is needed to obtain a reliable average. Obtaining average times for relatively rare case events may require a separate sample of only those events be taken.

Research should be done to establish required time frames and policies on continuances.

TYPICAL SEQUENCE OF HEARINGS FOR FOSTER CHILD UNABLE TO RETURN HOME



STEP 4: CALCULATE AVERAGE TIME PER EVENT

After selecting the particular events to be included in the workload assessment, the next step is to determine the *average* time spent on each event type. This can be done with an actual time study or with estimates derived from a panel of judges working iteratively. We recommend a

process that measures actual average time for each event, later adjusted by a panel of experts to take into consideration the amount of time an event *ought* to take. It is even possible for this panel of experts to estimate the average time each event currently takes using this process, although we recommend grounding at least some of the estimates in reality by having judges keep a log of time spent or having observers time key events.

Judges and court support staff are asked to record the time it takes to process specific events.⁴⁵ The longer, discrete events should be tracked individually as separate entries on the time log. This is the most accurate way to record elapsed time, but it is also the most labor intensive and, therefore, more costly.

Repetitive, short case-processing events may be measured by volume. For example, if a judge held six shelter hearings in a two-hour period, the log form could simply record “six shelter care hearings two hours” and the analyst could record that the hearings averaged 20 minutes apiece. At a later point in the study, a Delphi-type analysis (this will be discussed in more detail in Step 10 below) may determine whether 20 minutes per shelter care hearing is sufficient to permit best practice and, if the answer is no, an adjustment may ultimately be needed to determine the task weight for shelter care hearings. For example, a Delphi-type analysis might conclude that judges require an average of 40 minutes to conduct a shelter care hearing that will address all issues called for by state law and best practice as called for by the *Resource Guidelines*,⁴⁶ including issuing and distributing legally sufficient orders at the end of each shelter care hearing.

⁴⁵ Observers, especially students, can be hired to time events that take place in the courtroom as one method of obtaining court time per event. Care should be taken in training students so that they can easily distinguish one case hearing event from another (for example, some courtrooms do not clearly and audibly note the purpose of each hearing before the hearing begins). Even so, judges must be involved in the study to record the time spent in chambers or in other off-the-bench activities.

⁴⁶ *Supra*, note 5.

⁴⁷ Care should be taken in the training and supervision of such students to ensure accuracy of case file coding.

⁴⁸ The average accounts for the fact that judges with particular expertise in some areas of law or skills at reaching accommodation may be able to dispose of certain types of cases more quickly than other judges.

STEP 5: DETERMINE FREQUENCY OF CASE-PROCESSING EVENTS

The purpose of this step is to determine how frequently key case-processing events occur in each of the case types being studied. Accurate information on the frequency of event occurrence is as critical to the success of a weighted caseload study as is accurate information on time. Closed cases should be selected for review so that each event has the possibility of being present.

In practice, accurate and reliable frequency of event information is one of the most difficult data elements to obtain. There are two primary sources for data on the frequency of case-processing events: 1) Automated Information Systems. If data on event frequency are collected in a consistent and accurate fashion, the automated system will be a fast, cost-effective source of critical information. 2) Sample of Case Files. Without an automated information system, staff must review case files one at a time. This task can be made easier if law students or graduate students can be hired and thoroughly trained to review case records.⁴⁷

STEP 6: CALCULATE THE TASK WEIGHTS

If all case-processing events occurred in every case, then event weights could merely be added to calculate case weights. However, not all cases require all events to be resolved. Consequently, event time must be weighted by the frequency with which it occurs in the typical case to construct the case weights. A “task weight” involves multiplying the time an event takes by the frequency with which it occurs.

For example, suppose that a shelter care hearing occurs in 93 percent of the cases and the average estimated duration of each hearing is 15.8 minutes. In contrast, an adoption hearing occurs in 15 percent of the cases and takes an average of 20 minutes (that average includes both contested and uncontested adoption hearings). The *task weight* for a pretrial detention hearing is then 14.69 minutes (15.8 minutes x .93) and for an adoption hearing is 3 minutes (20 minutes x .15).

The duration of each event multiplied by the frequency of its occurrence produces a task weight.

Applying a Delphi analysis, suppose that experts found that shelter care hearings should average 40 minutes each to reflect “good practice” and that 20 minutes is an appropriate average time for an adoption hearing. The task weight for a shelter care hearing would then be 37.2 (40 minutes x .93) and the task weight for an adoption hearing would remain 3 minutes (20 minutes x 0.15).

STEP 7: SUM TASK WEIGHTS TO PRODUCE THE CASE WEIGHT

Summing the individual task weights results in a final case weight, which is simply the amount of time spent on a typical case. A case weight of 60 minutes means that, on average, those types of cases require 60 minutes to resolve. It is important to emphasize that these case weights are an average, and some cases will take much longer to process while others will be disposed more quickly.⁴⁸ States using a panel of judges to estimate case

weights tend to have higher weights for abuse and neglect cases. For example, West Virginia judges estimated that abuse and neglect cases require 585 minutes to resolve. California, using an event-based Delphi process, estimated 397 minutes per dependency case, and Florida estimated the average dependency case to take 281 minutes.⁴⁹ States measuring actual times to resolution have lower average weights (e.g., Michigan 271 minutes, North Dakota 172 minutes, and South Dakota 220).⁵⁰

STEP 8: DETERMINE JUDGE TIME AVAILABLE TO PROCESS CASES

Once the time necessary to dispose of different types of cases has been determined, the next question is “How much judge time is available to process cases?” The time calculations are in two parts – number of days per year available to judges to process cases and the number of hours per day available for case-related work. Multiplying these two measures together gives the “judge or clerk year” – the amount of time the “average” judge or clerk has to process cases during the year case-related activities both in court and in chambers. The judge or clerk year is an estimate of the average amount of time a judge has available to process both case-related and non-case-related work.

WORKDAYS PER YEAR

In establishing the “average” or “standard” judge year, the number of days available to hear cases must be reduced by subtracting weekends and holidays, as well as time related to illness, vacation, and judicial education. Some

judges also invest several days working with child welfare and treatment agencies to better serve the needs of children. The court then becomes part of the community team⁵¹ helping to create partnerships to advocate for new services and resources for children.

The judge year calculation is not straightforward because judges are not allotted a set amount of days for vacation and illness, or even told how long a day they should work, as are other state employees. Instead the number of days must be estimated, but most states have similar numbers of workdays per year.

Kansas	224	Michigan	215
Missouri	224	New Mexico	214
Delaware	222	Washington	214
New York	221	Connecticut	213
Colorado	220	Wisconsin	213
Georgia	220	Nebraska	211
Oregon	220	Utah	211
Rhode Island	220	Louisiana	209
Arkansas	218	West Virginia	209
Hawaii	218	North Dakota	205
South Dakota	216	Minnesota	202
Florida	215	Alabama	200
California	215		
			25 State Avg. 215

⁴⁹ Victor E. Flango et al., “West Virginia Redistricting Study,” unpublished report to the Administrative Office of State Courts, October 15, 1998; Brian J. Ostrom et al., “California Judicial Workload Assessment,” unpublished final report, August 2001; Brian Ostrom, *Florida Delphi-based Weighted Caseload Project: Final Report* (Williamsburg, VA: National Center for State Courts, 2000).

⁵⁰ Brian Ostrom and Neal Kauder (eds.), *Examining the Work of State Courts, 1997: A National Perspective from the Court Statistics Project* (Williamsburg, VA: National Center for State Courts, 1998), pp. 94-95.

⁵¹ Jackson County, Oregon, uses the term *community family court* to represent the commitment to partnerships among the court, community, and service providers. See First Judicial District: Jackson County, “Family Advisory Committee Findings and Recommendations,” December 1998, p.6.

⁵² Judges or courts that spend inordinate amounts of time on administrative matters should, as a general rule, be excluded from any formula used to calculate the judge day. However, accounting for these more extreme cases can be important in determining judgeship needs on a more local or individual level. Adjustments for these types of cases can be made *normatively*, that is, after the initial weights or judgeship needs have been assessed empirically. In this event, a series of policy decisions concerning special or additional judgeship needs must be made.

LENGTH OF JUDGE DAY

The judge day is separated into two parts: the amount of judge time devoted to (a) case-related matters and (b) non-case-related matters. A judge may work a nine-hour day, but only part of the day is devoted to hearing cases.⁵²

Although judicial time available to process cases will vary daily, the *typical* day will include the number of hours in the workday minus deductions for:

- administrative time, correspondence, phone calls
- travel time
- docket management
- waiting or “dead” time
- public education
- civic activities
- general and legal research
- opinion writing
- collaborative meetings

For example, starting with an average 9-hour workday, and by extension, total available time of 116,100 minutes (9 hours x 215 days x 60 minutes), *case-related* time is calculated by subtracting 1 hour for lunch and 2 hours of administrative time (including travel time). The 9-hour day does *not* take into account judges who work extra evening hours because of crowded dockets or additional responsibilities related to best practices in child abuse and neglect cases (e.g., convening collaborative meetings around systems’ reform).

JUDGE TIME

The judge year is then calculated by multiplying the number of judge days available by the number of case-related hours in the day. Judge years are not strictly comparable between states because some states include the time necessary for legal research, typically a non-bench activity, in a judge day, while others build these activities into the individual case weights. Therefore, case weights for similar types of cases may be smaller in some states than others depending on how non-bench activities are counted.

However, it is appropriate to adjust the judge year in dependency cases to take into account dependency judges’ need to be involved with their communities. For example, to prevent needless court delays, many dependency judges must periodically meet with child welfare administrators to ensure that those agency employees who frequently serve as witnesses consistently appear in court and appear on time and that they come prepared. Similarly, many judges must meet with agency administrators and their attorneys routinely to ensure that their petitions and court orders are complete and meet legal requirements.

More generally, it usually is not enough to take into account the time that dependency judges currently spend on these and other off-the-bench activities. Judges currently may not take enough time for training and community involvement as is necessary to represent best practice. The judge year, like case events, may need to be adjusted to permit sufficient time for non-case-related activities for child abuse and neglect cases.

STEP 9: RELATE CASE WEIGHTS TO JUDGESHIP NEEDS

Once total filings by case type are compiled for a given year, they can be multiplied by their respective case weights. Summing these “weighted filings” provides an estimate of the total amount of judge time required to process a given annual caseload.

The primary purpose of case weights is to determine the need for judges in courts throughout the state. The use of case weights as the basis for estimating judicial need in any given county depends on the weights being applied to a large number of cases – typically a year’s worth of filings. Indeed, estimates of judgeship needs should be made on the basis of *projected* filings, to partially accommodate the time lapse between when the request for judgeships is made and when judgeships are actually authorized.

STEP 10: ADJUST THE CASE WEIGHTS TO ALLOW FOR ADDITIONAL FACTORS

It is important to remember that even the most widely used and accepted judicial workload assessment will not objectively determine the exact number of judges needed to stay current with caseloads. No strictly quantitative resource assessment model by itself can accomplish that goal. Instead, a quantitative model can only approximate the need for judicial resources and provide a benchmark for comparison among judicial jurisdictions. The results can then be used in concert with other considerations, including budget constraints, population trends, and other more qualitative, court-specific factors that may

differentially affect the need for judicial resources statewide.

No set of statistical criteria will be so complete that it encompasses all contingencies. Each court jurisdiction will have peculiarities in caseload caused by differences in demographics and other factors. Administrative responsibilities of the judges, variations in case-processing practices, and the need to coordinate with child welfare and community treatment agencies are legitimate mitigating factors in judgeships and court support staff allocation decisions. Distance from the nearest court may be another qualitative consideration – what is a reasonable distance from courts for the public? A decision to add a judgeship in a rural area that would save the populace from driving for two hours may justify a particular judgeship before quantitative measures, such as case filings, show that one is needed. This consideration points to the need to consider access to justice as a qualitative consideration for rural courts.

ADJUSTING CASE WEIGHTS TO ALLOW FOR MEASURES OF BEST PRACTICE

One recurring criticism of workload assessments based on weighted caseloads is that they simply codify current practice. Judges may spend a total of 210 minutes over the life of an average child abuse or neglect case because that is what is required to make a fair decision or because that is all the time they have available to devote to that case and still process the other cases on their docket. One way to reassure ourselves that the time devoted to each case and each case event is adequate is to derive case weights from only the most productive and well-

...even the most widely used and accepted judicial workload assessment will not objectively determine the exact number of judges needed to stay current with caseloads.

⁵³ See the discussion above regarding the complexity of identifying courts with short processing times in child abuse and neglect cases. Another consideration in choosing such courts might be whether they exceed national standards established by the federal government for case completion. For example, a national standard adopted by the U.S. Children's Bureau, Department of Health and Human Services, specifies that 32 percent or more of the adoptions of children in foster care should occur within 24 months after children were removed from home. Selections of courts might be based, in part, on the percentage of their adoptions of foster children that occur within 24 months after children's placement into foster care.

⁵⁴ To "fully address" these issues requires the following: bringing up facts that are specific to the particular case, allowing all parties and counsel to speak about the issues, and making sure that all parties understand what has been said and decided about the issues.

managed courts (those whose case-processing times are shorter than the statewide average) in the state.⁵³

Another method of adjustment takes into account the opinion of experts, in this case judges, to review the actual time taken to resolve cases as well as the time it ought to take. How long should case events take in order to give a judge sufficient time to ensure full and fair consideration of all matters relevant to case determination? Is a total of three and one half hours spent over the life of an average dependency case too long or not long enough?

In this step of the process, judges are provided with the objective measures of the time it is currently taking to resolve cases and then the judges are asked if any case events deserved more time than they were currently being allotted. For example, assume an uncontested shelter care hearing is currently averaging 15 minutes, meaning in practice that the average hearing takes between 10 and 20 minutes. Judges could be asked if that is a sufficient amount of time to "fully address" all of the issues,⁵⁴ such as:

- Reasonable efforts to prevent placement
- Whether removal from home is in the child's best interests
- Whether parent wants child's return home
- Location and notification of missing parties
- Current service needs of children (mental, medical, educational, etc.)

- Appropriateness of current placement
- Visitation with parents, if child is placed away from home
- Visitation with siblings, if children are separated
- Need to place child away from custodial parent
- Need to place child away from noncustodial parent
- Location of relatives and appropriateness of placement of child with relative other than parent

And whether they follow best practices:

- Prepare and distribute court orders in uncontested hearings before the parties leave the courthouse
- Schedule the next hearing
- Introduce the persons present
- Explain the hearing
- Advise parties of their rights at the hearing
- Explain case deadlines and the consequences of missing deadlines.
- Make case-specific findings regarding reasonable efforts
- Make case-specific findings on the need to remove the child from home

Based on the judges' answers, they might adjust the time needed for shelter care and other hearings. For example, assume, as above, that the average time for shelter hearings is 15.8 minutes but to achieve identified best practices associated with this hearing, the time for a shelter care hearing needs to be 40 minutes. Also assume, as above, that the frequency of shelter care hearings is .93 per case. The task weight for shelter care hearings would increase from 14.7 (15.8 x .93) to 37.2 (40 x .93) to accommodate time needed for best practices. With adjustments for other hearing types, the total case weight for dependency cases would also be modified.

This example shows how time study weights can be adjusted to take into account aspirational or best practice goals. How much time should it really take to conduct a shelter care hearing? The *Resource Guidelines*⁵⁵ for child abuse and neglect cases recommends 60 minutes to make a preliminary protective or shelter care hearing “as thorough and meaningful as possible” and to hear from all interested parties present.⁵⁶ They also suggest that a noncontested adjudication hearing should take 30 minutes as should a disposition hearing and a six-month review. A combined adjudication and disposition hearing should take 60 minutes, as should a termination of parental rights hearing.

National guidelines, however, may not be an exact fit for a particular jurisdiction. For example, some courts hold meetings outside the presence of the judge in which they discuss, narrow, and define issues for shelter care hearings. While the judge would still need to review the issues for the hearing and ensure that all parties understand the

court's decision, the pre-hearing meeting may eliminate the need for some of the in-court discussion. In addition, the length of a hearing may vary, in part, depending upon the amount of pre-hearing preparation. To the extent that the court is successful in communicating its expectations, attorneys and representatives of child welfare and treatment agencies may present better written court reports and verbal testimony, thus reducing the amount of time the judge needs to question the parties. On the other hand, if there is poor case preparation, “task weights” may require upward adjustment. Another limitation of the times specified in the *Resource Guidelines*⁵⁷ is that they do not include contested cases. For example, if a certain percentage of shelter care hearings become contested hearings involving the examination and cross-examination of subpoenaed witnesses, the frequency and average length of *contested* shelter care hearings needs to be taken into account.

Knowing the actual amount of time currently being spent on case processing allows judges to explicitly decide what essential elements ought to be included in each event and to estimate the time needed to conduct that event. Objectively derived calculations of task time help to keep the discussion realistic. In other words, if an event is currently taking 15 minutes and judges are all fully engaged in all key issues and are able to take such steps as completing and distributing court orders on the spot, it may be unrealistic to think that judges would now spend an hour on that particular event. Having that discussion, however, is important when explicitly deciding what steps in the procedure are essential and cannot be compromised and which aspects of a case event may be

⁵⁵ *Supra*, note 5.

⁵⁶ *Ibid*.

⁵⁷ *Ibid*.

The *Resource Guidelines* for child abuse and neglect cases recommends 60 minutes to make a preliminary protective or shelter care hearing...

⁵⁸ This is an intermediate step between making estimates based upon case type alone and using all case events, but is not dependent on case type.

shortened by prior preparation. This discussion is also necessary to the principle that all cases coming before the court be treated similarly.

An additional example may be useful to illustrate how an adjustment might be made. To simplify the illustration, let us consider only three case events: pre-hearing activities, hearing (adjudication/disposition/termination of parental rights), and post-hearing reviews.⁵⁸ Let us further assume that empirical research has shown that in child abuse and neglect cases, pre-hearing activities occur in all (100 percent) of the cases, but that only 63 percent require major hearings, and 70 percent require post-hearing reviews. We also know that pre-hearing activities average 68 minutes, all major hearings combined average 87 minutes, and post-hearing reviews average 69 minutes, and that a judge has 6 hours per day to devote to processing specific cases (as opposed to handling docket administration, dealing with the public, doing research, and performing all other judicial activities not directly related to resolving cases).

The case weight for an average child abuse or neglect case in this jurisdiction would then be estimated to be 171 minutes (summing the three task weights: 100 percent of 68 minutes of pre-hearing activity, 63 percent of the average hearing time of 87 minutes, and 70 percent of post-hearing reviews averaging 69 minutes) and a judge could be expected to resolve 452 child abuse or neglect cases in a year.

At this point, the usual time study would be complete. But what if an expert panel of judges wanted to examine alternative assumptions? For example, what would

Adjusting Case Weights	
Occurrence Rates	
Pre-hearing activities	100%
Major hearing	63%
Post-hearing reviews	70%
Time in Minutes	
Pre-hearing activities	68 min
Major hearing	87 min
Post-hearing reviews	69 min
Case Weight	171 min
Filings per Judge per Year	452 cases

happen if judges decided that *all* cases required some post-hearing review (so that the post-hearing occurrence rate increased from 70 percent to 100 percent) and that a few more cases required a hearing (increase the hearing rate slightly, from 63 percent to 65 percent)? What would be the impact of those changes on judicial resource needs? Assuming no changes in the measured time each event took and no change in the judge workday or year (215 days), these two changes would increase the case weight from 171 minutes to 194 minutes, which means that the number of child abuse and neglect cases a full-time judge could handle in a typical year would decrease from 452 cases to 399 cases. In many circumstances, a change of this magnitude would not be sufficient to require an additional judge. For example, if this jurisdiction typically had 800 dependency filings per year and had 2 judges, the decrease in workload would not justify the addition of a judge.

What happens, however, if the changes suggested above are made and judges decide that three post-hearing reviews are required, rather than one, and that each should take about one hour? Changing that one assumption, the average time required to process each case now increases hearing time to 180 minutes (rather than the 69 minutes the one hearing now takes). This change in the number of hearings increases the case-processing time from 194 minutes to 305 minutes, and consequently reduces the number of cases that a judge could handle in a year from 399 to 254. A change of that magnitude in a court that hears 800 dependency cases per year would require the addition of a new judge.

The panel could adjust for other factors, as well, for example, increased time taken in pre-hearing activities, longer hearings, or reduced time in a judge year to determine the impact on judges needed to keep current with cases. With all of this information before them, the steering committee may better decide which solutions are feasible without the addition of judges, or how best to assign a new judge to make the maximum impact on workloads.

STEP 11: COMPARE JUDGES NEEDED TO ACTUAL NUMBER OF JUDGES BY JURISDICTION

This will show which jurisdictions have the greatest *relative* need for judges and court support staff.

Adjusting Case Weights Part 2			
		1st	2nd
Occurrence Rates	Time Study	Adjustment	Adjustment
Pre-hearing activities	100%	100%	100%
Major hearing	63%	65%	65%
Post-hearing reviews	70%	100%	100%
Time in Minutes			
Pre-hearing activities	68 min	68 min	68 min
Major hearing	87 min	87 min	87 min
Post-hearing reviews	69 min	69 min	180 min
Case Weight			
Filings per Judge per Year	171 min 452 cases	194 min 399 cases	305 min 254 cases

STEP 12: HAVE STEERING COMMITTEE REVIEW FINAL RECOMMENDATIONS AND PROVIDE FOR UPDATES

Statistical models cannot anticipate all possible contingencies. It may not even be desirable to do so, because then the established judgeship needs criteria would have to be nearly as complex as the real world. Resource need models that require a large number of data elements are more costly to produce than is justified by the added precision, not to mention the problems associated with ensuring accuracy of the data.

AN INDEPENDENT PERSPECTIVE IS NECESSARY

An independent review to determine whether a court appearing to need judges could reduce that need through

⁵⁹ The Delphi technique was first developed by the Rand corporation in 1964 (O. Helmer, "Convergence of Expert Consensus Through Feedback," Rand Corporation, 1964) and applied to courts in Michigan by David P. Doanes, "The Effect of Case Weights on Perceived Court Workload," *Justice System Journal* 2 (spring 1977): 270.

Unlike a focus group, the Delphi process is iterative – it is a search for consensus.

operational changes remains a good idea. Quantitative data should be tempered by experience. Any jurisdiction should be able to request an evaluation of current operating procedures and have the ability to present a case for adding a new position. The procedure should include a method for systematically soliciting local opinion. A team of qualified individuals should be selected to visit each site to determine if present resources could be used more efficiently and if new procedures or practices would lead to less demand for resources. These site visit teams should consist of judges, court administrators, and court clerks from local or national jurisdictions that (a) have a reputation for good management and efficient case processing and (b) thoroughly understand best practices in dependency litigation. Team members should not be chosen from circuits or districts that are immediately adjacent to the circuit or district making the request to avoid the appearance of a conflict of interest. The on-site visit will benefit not only those challenging the guidelines, but also the team members, who will be exposed to alternative procedures. Indeed, the visit could stimulate change in the home courts of team members.

Courts that desire new judgeships or court support staff should be able to demonstrate their need despite the implementation of administrative and procedural changes designed to reduce or avoid the need for new judgeships. A review could also determine whether the need is long-term or is caused by a temporary increase in filings or unusually difficult dispositions.

Quantitative criteria for assessing the need for judges and court support staff not only become inflexible if not

tempered by careful consideration of local conditions and best practices, but also become less of a guide if they are modified or waived with each challenge. The quantitative criteria may need to be reexamined if examination of qualitative criteria results in frequent changes.

THE DELPHI APPROACH⁵⁹

This chapter presented a 12-step method for estimating the time necessary to process court cases. The Delphi approach can be used in two steps of that process: 1) as a method of estimating the time necessary to process cases and 2) as a method of adjusting case weights already obtained from a time study.

WHAT IS DELPHI?

Delphi is technique that can be used to obtain estimates from a panel of experts. It differs from a focus group in that the experts have the opportunity to change their own estimates based upon feedback from other panel members. Unlike a focus group, the Delphi process is iterative – it is a search for consensus. According to McDonald and Kirsch:

Briefly stated, the Delphi Method of case weighting is a way of developing case weights by single estimates generated by a panel of experts. Initially, experts are asked to estimate the amount of time that they believe is necessary to dispose of various types of cases in their jurisdictions. Their responses are then averaged and this average is shown to them in a second round of questioning. Experts will either adhere to their initial response or modify them to more closely approximate the group average. This

process can be repeated until a group consensus emerges as to the average amount of time spent disposing of each of the various types of cases.⁶⁰

USING DELPHI TO ESTIMATE CASE-PROCESSING EVENTS OR TOTAL CASE WEIGHTS

The Delphi process can be used to estimate the average time required to dispose of each case event or may be used to estimate the time required to dispose of the case as a whole.

ESTIMATING CASE EVENTS

Estimating the time necessary to process key events in a case is very similar to measuring the time required to process key events. The difference is that rather than measure the actual time it takes to complete an event, such as a shelter care hearing, judges are asked to estimate the time needed to conduct a shelter care hearing. Sometimes judges may be asked to assume that a case event, such as termination of parental rights hearing, is either contested or uncontested so that their estimates can be more precise.

These estimates are then compiled and each judge on the panel receives his or her original estimate together with the average estimate of the time needed to process each case event as determined by the group. Each judge then has the option to revise his or her initial estimate, and new averages are calculated until consensus is reached.⁶¹ Using dependency cases as an example, the team would identify:

1. Key events, on the bench and in chambers, involved in a child abuse and neglect case⁶²
2. How much time it takes to complete each event

ESTIMATING CASE WEIGHTS

Sometimes, rather than go through many case events, Delphi process leaders simply have judges estimate the time necessary to process a specific case type, for example, dependency, delinquency, or divorce. Estimating case weights directly may be less precise than measuring case events, but is less time-consuming because there is no need to compile data on specific case events and to compute task weights. On the other hand, case weights are sometimes more accurate in the aggregate, because there is only one estimate. A small overestimate in as many as 49 separate case events sometimes leads to a large case overestimate. Although there is not sufficient information to draw a definite conclusion on these two processes, a Delphi process used in Florida resulted in an estimate for the average dependency case to be 281 minutes.⁶³ In contrast, a Delphi process used in a California site based on a series of case events found the average dependency case to be 397 minutes. (See discussion under *Step 7: Sum Task Weights to Produce the Case Weight* above.)

Other than separating the case into separate events, the rest of the Delphi process for estimating case weights as a whole is the same as that described above. Case weights are estimated and averages are calculated, and returned to each judge with a request to adjust their original estimates in light of the information provided by their

⁶⁰ H. Graham McDonald and Clifford P. Kirsch, "Use of Delphi Method as Means of Assessing Judicial Manpower Needs," *Justice System Journal* 3 (spring 1978): 314.

⁶¹ The standard deviation is a way to measure the amount of agreement among judges on the various rounds of questionnaires to determine if panelists are approaching consensus.

⁶² Estimates used in Delphi typically use fewer case events than the traditional time study. Sometimes, cases are only divided into three parts: pre-hearing (pretrial) activities, major hearings (trials), and post-hearing (post-adjudication) activities.

⁶³ Brian Ostrom, *Florida Delphi-based Weighted Caseload Project: Final Report* (Williamsburg, VA: National Center for State Courts, 2000).

⁶⁴ *Supra*, note 5.

⁶⁵ *Ibid.*

colleagues. After the first iteration, some states calculate ranges (either medians and percentiles or means and standard deviations) and use these ranges as guides for the second iteration of the questionnaire. Note that even with closed-ended questions, judges may select times outside the set ranges. The process is repeated until consensus is achieved or until the range of responses is reduced as much as it can be.

USING DELPHI TO MAKE QUALITATIVE ADJUSTMENTS TO CASE WEIGHTS

In addition to estimating the time necessary to resolve case events or types of cases, the Delphi process is useful to estimate the amount of time a particular hearing *should* take to resolve if the hearing is to provide sufficient time to hear from all parties and to follow best practices.

Asking judges to specify the amount of time necessary to conduct a proper hearing, regardless of the amount of actual time the current hearings in fact take, provides judges with the opportunity to review their performance and to consider what might be accomplished if more time were available. Indeed, the encouragement of discussion among judges as to what practices are best is a benefit in itself in that judges can come to consensus on what procedures are absolutely essential to a fair hearing and which are optional. This consensus helps to ensure fairness because with judicial assumptions made explicit, judges are more likely to employ consistent procedures. A refinement of this approach, which we recommend, is to first seek *consensus* on what steps should typically occur in a particular court event, before estimating how much time it should take. To begin this analysis, a list of

steps associated with a court event might be derived from a combination of (a) what state statutes explicitly require, (b) what, if anything federal law requires, and (c) nationally accepted best practice guidelines such as those set forth in the *Resource Guidelines*.⁶⁴ For example, the following are steps involved in a shelter care hearing (also referred to as a “preliminary protective hearing”), according to federal law and the *Resource Guidelines*.⁶⁵ (See *Packard Court Performance and Judicial Workload Assessment Toolkit* for similar lists of events for each hearing type.)

- Prepare and distribute court orders in uncontested hearings before the parties leave the courthouse.
- Introduce the persons present.
- Explain the hearing to parties.
- Advise parties of their rights at the hearing.
- Explain case deadlines and the consequences of missing the deadlines.
- Make case-specific findings regarding reasonable efforts.
- Make case-specific findings on the need to remove the child from home.

- Ask questions or receive evidence from the parties on each of the following issues:
 - Reasonable efforts to prevent placement.
 - Whether removal from home is in the child's best interests.
 - Whether parent wants child's return home.
 - Location and notification of missing parties.
 - Current service needs of children (mental, medical, educational, etc.).
 - Appropriateness of current placement.
 - Visitation with parents, if child is placed out-of-home.
 - Visitation with siblings, if children are separated.
 - Need to place child away from custodial parent.
 - Need to place child away from noncustodial parent.
 - If placement with parents is not practical, efforts to locate relatives and determine the appropriateness of placing children with relatives if placement other than parents.

So, for example, beginning with the above list, supplemented by any additional requirements of state law, the Delphi group might be asked whether this is an

appropriate set of steps to complete during a typical shelter hearing. If not, which steps should be added or subtracted? After the discussion on this issue is complete and the Delphi group has reached a consensus on what steps should be occurring during the shelter hearing, the group is asked to estimate how long it would take to conduct a typical shelter hearing if each of the recommended steps is completed.

Again, estimates made in isolation tend to be high, and it can be expected that judges will tend to report needing much more time for each event. Having available the *actual* times associated with each case event grounds this exercise in reality, and helps judges decide in context which case events most need the additional time. Having the actual, measured time per event allows judges to explicitly decide what essential elements ought to be included in each event and to estimate the time needed to conduct that event. The objective numbers help to keep the discussion realistic. In other words, if an event is currently taking minutes, it may be unrealistic to think that judges would now spend an hour on that particular event. It may not be unrealistic, however, to devote 45 minutes to that event at the expense of other, less critical events.

In brief, the Delphi process allows judges to ask "what if" questions and that helps them make many practice decisions, including what could be done with additional resources. What would happen if judges increased review hearings from 15 minutes to 30 minutes? What would be the result on resources if a review hearing were scheduled every 90 days rather than every 180 days? What would be the result if judges spent three additional days leading

community discussions rather than on the bench? How would the picture change if a quasi-judicial officer, for example, a magistrate or commissioner, were added to the court, rather than a judge? Considering these alternatives is likely to improve the employment of scarce judicial resources.

BENEFITS AND LIMITATIONS OF USING THE DELPHI METHOD

Delphi is used to estimate case weights and case event time because it is much less costly than measuring the time necessary to complete each event. In addition, the Delphi approach provides an estimate of how long events would take if they were modified to follow best practices.

Delphi is a way to substitute opinion for observation and has the advantage of involving participants in the process of workload assessment. Having judges participate in the creation of case weights gives the weights more credibility because judges know how they were derived.

Consequently, the final case weights are more likely to be accepted.

The reliability and validity of the estimates are always constrained by the depth and breadth of experience of the opinion holder. Even when judges reach consensus on the times required to process various case events or case types, the estimates may differ significantly from actual processing time as recorded by impartial observers, or even the times recorded by judges themselves.

An iterative process is used to estimate both case weights as a whole or case events because experience has shown that the initial tendency of many judges is to over-estimate the time it actually takes to process a case. This is not unexpected since judges tend to remember the longer, more complex cases. The most memorable cases will be the ones that stand out from the rest due to an extra measure of contentiousness, unusual or interesting issues, or frequency of hearings and duration of hearings. The cases become “anchors” in the estimation process. It is apparent that all of these characteristics will be more common to cases that require more judge time than the average case. On the other hand, many cases flow through the court that require very little judicial involvement, perhaps just signing orders of dismissal, so that judges may find it difficult to include these cases in their estimates. If the court does use judicial officers, they should be included in the Delphi process.

Various strategies can be used to improve the time estimates provided using the Delphi method, including:

- Involving many judges, perhaps even all judges or a large representative sample, in the processes so that the expert opinion is increased.
- Cross-checking estimates against timed measures for particular case events or case types to keep the process grounded in reality. For those events in which there is much disagreement, ask judges and support staff to record the amount of time it takes to complete that case event for all of the cases they hear in the next couple of months. For example, judges should not record the time it takes to complete *all* hearings, but

only selected hearing where estimates vary the most, such as termination of parental rights hearings.

- Triangulation – using additional data collection procedures and sources of data to supplement the Delphi study.



While judges typically overestimate the length of time hearings take (when compared to reality), they typically underestimate the length of time when asked to consider how long routine child welfare hearings *should* take. This underestimation might especially occur in child welfare cases because of the following factors:

- High workload demands make it practically impossible for judges to fully comply with the expanding legal requirements in child welfare cases or to perform best practices – such as the steps listed above for shelter care hearings. In response to this reality, judges may conduct truncated hearings. This reality may then affect the estimates provided when judges are asked to estimate how long hearings *should* take.

- Because child welfare law and practice used to be less complex than it is today, court staff have become used to setting aside short times on the court docket for most types of child welfare hearings. Judges may have learned to do what they have to do to conduct their hearings within the short times set aside on the calendar.
- Attorneys handling other types of litigation demand more time for those other case types, and partly in response to this pressure, judges may keep child protection hearings short.
- Court forms may not call for detailed, individualized findings. Court forms may be kept short by judges who feel that they do not have sufficient time for best practices (i.e., detailed, individualized findings) in child welfare cases.

The *Packard Court Performance and Judicial Workload Assessment Toolkit* contain examples of Delphi forms and instructions, including a set of sample instructions for obtaining Delphi case weights by mail.

CONCLUSIONS

Performance and workload measures when used together establish a baseline from which improvements in the processing of child abuse and neglect cases can be measured. They determine the level of performance achieved with current staff, and more importantly, what could be achieved with additional resources. What could be achieved includes not only quantitative improvements in the number of cases processed, but also measures of the **quality** of performance.

The major cost components in updating a workload assessment involve collecting event time and frequency information.

The number of abuse and neglect cases does not adequately reflect the work necessary to process them. These cases require more social services, more collaboration between courts and child welfare agencies, and more community involvement than most other types of cases. It is therefore important when examining the performance of any court handling dependency cases to include *both* the amount of time the judge must spend on case-related activities as well as the time spent on administration and collaboration, for example, speaking engagements at local bar association groups and schools; participating in meetings with other judges and judicial officers and in committee and group meetings related to court cases; attending system improvement and public outreach meetings; and attending judicial training programs. Finally, the dependency court judge's workload is tremendously influenced by the performance of people and entities providing information and resources to the court. Lack of time and resources on the part of social workers and attorneys can result in delays in the courtroom and ultimately affect the timely conclusion of dependency court cases.

Availability of workload measures will allow courts to examine the relationship between workload and performance and to determine the combination of resources required for judges to make more timely and thorough decisions. Objective measures enable courts to make a better case to legislatures or other funding bodies to provide the resources necessary to process abuse and neglect cases, and to demonstrate the impact of lack of resources on specific areas of performance. Resources alone are not expected to guarantee good court performance in terms of outcomes for children and

families, but the lack of adequate staffing will certainly hamper performance.

The workload assessment process described in this chapter illustrates an objective process for establishing a reasonable workload. Adding quality adjustments to the time study enables us not only to measure the amount of time currently consumed in processing abuse and neglect cases, but also the amount of time that it would take if resources were not a problem and judges could fully address each issue involved in the cases. The hybrid process of using time studies with quality adjustments permit us to answer the question of what improvements would be possible if more judges and court support staff were available.

To be useful, case weights need to be reviewed by a steering committee of local experts and must be periodically adjusted and updated to ensure that they continue to accurately represent workload. The credibility of weights suffers if they become obsolete. This periodic updating is necessary to reflect changes in case-processing event times that may result from increased efficiency, statutory changes, or case management initiatives. The major cost components in updating a workload assessment involve collecting event time and frequency information. One way to reduce this cost is to monitor samples of those case events that exert a strong influence on the overall case weight (events that take a large amount of time or occur with high frequency). These samples need not be large enough to accurately establish case weights, but rather to serve as trip wires to alert court officials of possible changes in event times. If specific "high-impact" case event times appear to be

changing, it may be worthwhile to take a statistically valid sample of the questionable event to determine whether the task weight should be adjusted. To the extent that automated systems can be modified to collect data on judge time per event or specialized dockets permit the calculation of elapsed time for specific case events, data collection will be easier.

In sum, the recommendation is to sample case types and case events periodically from a rotating sample of court jurisdictions, and then to update only those case event times within specific case types that have changed significantly. In other words, it is not necessary to continuously monitor all case events for all case types. Time and money can be saved if a credible review process is instituted that tracks only the most critical case events and corresponding case weights.

Chapter 4

The Planning Process – How Your Court Can Make It Happen

INTRODUCTION

How *can* your court make court performance measurement and judicial workload assessment happen? What are the elements, steps, and procedures courts should implement when embarking upon current performance studies and judicial workload assessments? It begins with effective project planning, proceeds with effective project implementation, and ends with effective project completion and next-step strategizing.

In a nutshell, effective project planning includes deciding on the objectives of the study, its focus, and the overall scope of the project. It involves obtaining a commitment to the effort and the use of the study's results – ***obtaining a consensus about the purpose and nature of the study.*** It involves appointing a project or steering committee and designing and implementing evaluation procedures. Anticipating the need for specific and overall project monitoring and project evaluation should also be part of project planning.

Planning can also include a basic analysis of the court performance measurement and judicial workload assessment process itself – a court might want to do this before attempting any study of this magnitude to predict potential problems and then modify the overall project plan if necessary.

There are three phases to any project: (1) planning; (2) implementation; and (3) wrap-up.

And within these phases there are essential elements:

- Planning includes pre-project start-up, project design, and project administration
- Implementation includes actual data collection, oversight, and analysis
- Wrap-up includes generation of the report, dissemination, and next steps

PHASE 1: PLANNING

PRE-PROJECT START-UP

The first component of an effective overall project plan is determining the focus and scope of the study. The

⁶⁶ A note about high-profile events and cases – Any assessment of workload and performance must be attuned to the potential impact of events or cases that may have concentrated attention on court operations, compelled an examination of processes and actions, or otherwise focused interest on the court or social service agency in this area. Often, an especially tragic or otherwise high-profile case will alter system stakeholders’ behaviors and perceptions, and perhaps their willingness to participate or be candid in the proposed research. More importantly, such a case may mark a turning point for the court or the child welfare agency, one marked by dramatic changes in operations and performance. At that point, the timing of the assessment and time period for which cases or court records will be reviewed must be selected carefully, and any change in procedures or operations precipitated by the event should be duly noted in the analysis.

Packard *Guide* and *Toolkit* focus on two main areas of evaluation: (1) measuring the court’s performance in handling child abuse and neglect cases according to five core performance measures and (2) assessing the court’s judicial workload as part of that performance. Different courts will have different needs and areas of focus – we strongly recommend using *both* the core performance measures and the judicial workload assessment criteria as the principal elements of your court study – then adding specific areas of focus that may be of particular importance to your jurisdiction.⁶⁶ This not only will give the study needed structure, but also will help to ensure useful information that is relevant for your jurisdiction in planning for future improvement.

Gaining commitment for all phases of the study, including the use of the findings for strategic planning, is also a vital aspect of pre-project planning. Identifying who will do what tasks during the study and who will be responsible for oversight of the entire project at the beginning of the planning process will promote “buy-in” and shared responsibility for the study’s success. And part of gaining that commitment for all stages of the study depends heavily on developing and ensuring group ownership of the study’s process and outcomes. This requires that all key personnel involved – from leadership to project coordinators to data gatherers to consumers of the study’s report – understand and invest in why the court is doing the study and what it hopes to achieve from the results.

At the beginning of the court performance and judicial workload assessment project, make sure all the key players are involved in planning the project and that they remain involved during its evolution, outcome, and application.

Elements of gaining stakeholder “buy-in” include:

- Introducing key stakeholders to each other and the project by providing an overview of court performance and workload assessment approaches found in the preceding chapters of this *Guide*;
- Clarifying the reasons for the study and potential uses of study findings;
- Identifying and establishing a project advisory or steering committee by defining obligations and associated tasks of project participants; and
- Generating a list of project implementation challenges and solutions to overcome those challenges.

PROJECT DESIGN AND ADMINISTRATION

Project design is closely associated with, and should be the precursor to, the project’s overall administration – the importance of these components to the overall effectiveness of the project plan cannot be overemphasized. Unless your court’s study is clearly mapped out at the beginning of the project, and unless the project is administered effectively, deadlines will be missed, data will be lost, and results will be questionable at best.

Some of the key project design and administration elements that need to be mapped out at the beginning of your project, before data collection begins, are:

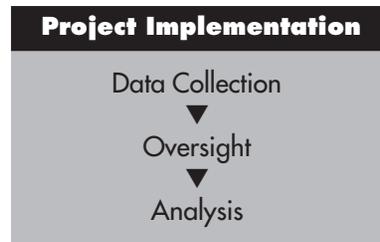
- Determining the types of data that need to be gathered, the best sources for these data, and the best methods for data collection and analyses.
 - Consider the core performance measures outlined in previous sections of this *Guide*, and the data elements needed to report outcomes on these measures. Discuss any outcome and process information that would be important to measure in your jurisdiction. Generate a list of possible data sources for these items and discuss data collection methods (see *Packard Court Performance and Judicial Workload Assessment Toolkit* for worksheets to assist with this process).
- Determining the accessibility of your data sources and the availability of the key or core performance measures (as well as any additional measures your court chooses to focus on).
 - Review your list of data sources and determine whether they will produce information on the required court performance measures. If not, what are some alternative plans for obtaining the measures (for example, case file review, interviews?) and what drawbacks might be associated with these alternative approaches (for example, more time intensive, more room for error, more qualitative?). Devise a plan to address any access or information-sharing concerns, or any concerns related to the source of the data.
- Determining what resources are available for data collection and analysis (e.g., staffing and time).
 - For example, is staff with the appropriate skills available to conduct case file reviews, interviews, or Delphi groups? Is staff with analysis skills available to the project?
- Articulating project monitoring and supervisory functions – who will train the data collectors and oversee data collection, analysis, and report writing to ensure consistency and continuity?
 - Who will have ultimate project oversight and accountability? Who will have day-to-day project-monitoring functions?
- Identifying a sample for study.
 - For example who will be your interview subjects or questionnaire recipients? What population of cases will you include from your analysis of the court’s management information system or review of case files? Who will be your “Delphi participants”? (See *Packard Court Performance and Judicial Workload Assessment Toolkit* for guidance regarding sample selection.)
- Identifying who the potential users of the results will be and what the best methods to communicate with them are.
 - For example, depending upon your audience or users of the study’s findings, would a report be more useful than a briefing? Or would both

communication methods be best? Would it be helpful to share interim or periodic progress reports with your audience before the final report is submitted? Would it be useful to share a draft final report for review and feedback before finalizing the report for more widespread dissemination?

PHASE 2: IMPLEMENTATION

DATA COLLECTION, OVERSIGHT, AND ANALYSIS

The second phase of the project primarily consists of data collection, oversight, and analysis. It is during this stage that the most “hands-on” activity will take place, as well as the most interaction with study participants (for example, judges, court personnel, and other stakeholders involved in the study). Data collection, oversight, and analysis must be consistent, coordinated, and conscious of the extraordinary demands on stakeholders’ time throughout the process.



Key components of this stage are:

- Training and supervising people to do the data collection.

- This includes assigning one or more persons to conduct periodic checks for data collection accuracy and data entry accuracy.
- Monitoring the project implementation and activities.
 - If project activities are constantly monitored, any problems that arise can be noted and midcourse corrections made to the implementation plan. In addition, any problems associated with the data collection procedures and instruments can be noted and adjustments made.
- Gathering the core court performance and judicial workload data (see previous sections of this *Guide*).
 - This may involve communication and coordination with other systems, such as the child welfare agency.
- Gathering the additional court performance and judicial workload data (if any).
 - As previously mentioned, courts may add measures to the list of core performance measures to address concerns specific to their jurisdictions. However, we would advise that any revisions to the instruments provided in this *Guide* be pretested to ensure that any additional data elements obtain the information they were intended to obtain. (See *Packard Court Performance and Judicial Workload Assessment Toolkit* for guidance on question construction and pre-test procedures.)

If project activities are constantly monitored, any problems that arise can be noted and midcourse corrections made to the implementation plan.

- Synthesizing all the data for analysis.
 - This may include transcribing interviews, questionnaires, focus groups, and Delphi responses; check-coding this information to ensure accuracy of coding; entering data obtained from case file reviews into spreadsheet programs for statistical analysis; check-coding this information for accuracy of data entry; and obtaining a management information system download where available.
- Analyzing, collating, and interpreting all the data. For quantitative studies it is appropriate to summarize information with indicators of central tendency, such as mean and median, and indicators of variability, such as standard deviation. Tests for statistical significance may also be performed and measures of relationship calculated (e.g., correlations). You do not have to perform statistical tests to produce a good report. However, doing so will make your conclusions stronger. Qualitative data obtained should be examined carefully because in addition to being a source of important information about context, history, opinions, and current practice, qualitative data help interpret, or provide explanations for, quantitative findings. (See *Packard Court Performance and Judicial Workload Assessment Toolkit* for guidance on how to analyze data.)

PHASE 3: WRAP-UP

GENERATION OF THE REPORT, DISSEMINATION, AND NEXT STEPS

The final stage of the project will be what your target audiences – including the public – will actually see. And the substance of this report or briefing will depend greatly on the quality and consistency of all the work you put into the planning and implementation stages above. In fact, politics aside, this final stage should be the most enjoyable and engaging part of the entire process – it’s your opportunity to provide valuable information to system constituents.

The report or briefing can also be used as the basis for further reports and updates on the court’s progress. Dissemination should be well thought out in advance – to ensure that key audiences are targeted, such as legislatures, bar associations, advocacy groups, and the judiciary (including appellate court judges), as well as court and system monitors, and the like. Finally, the “wrap-up” phase of the project can and should serve as the basis for defining next steps the court needs to consider in aspiring to achieve optimal performance.

The elements of this last stage are:

- Generating a report or briefing of your findings – always keeping in mind your target audiences.
 - Reports should address all project goals (including why some project goals were not met). Reports should address areas of importance to your target

Dissemination should be well thought out in advance – to ensure that key audiences are targeted, such as legislatures, bar associations, advocacy groups, and the judiciary...

audience. Reports should include, wherever possible, graphic representation of findings to facilitate understanding of findings. (See *Packard Court Performance and Judicial Workload Assessment Toolkit* for guidance on choosing a method of presentation and assembling your final report.)

- Sharing the report with an advisory or steering committee for review and feedback prior to any final dissemination and making modifications as appropriate.
 - This is a good opportunity to tap alternate sources of interpretation for findings. In addition, it helps to build “buy-in” for the final report.
- Disseminating the report or briefing.
 - Decisions should be made in advance about the communication strategy for the final report, including whether the report will be formal or informal, of if the format will be oral, written, or both. As previously mentioned, graphic representation of findings should be used wherever possible to facilitate understanding of the data.
- Convening a group to design a strategic or action plan that addresses study findings.
 - Findings can be used to target court improvement activities and to address judicial workload concerns, as well as other concerns specific to the court. Findings can be helpful to address controversial or political topics – often, “hard

data” can bring about movement on an issue that was previously a stumbling block to reform efforts. (See *Packard Court Performance and Judicial Workload Assessment Toolkit* for strategic planning worksheets.)

- Evaluating the quality of the project’s process and outcomes.
 - This process may include all of the above elements, but essentially determines whether the study has accomplished what it intended to accomplish. Were all of the project’s goals met or only some? What were the barriers to achieving project goals and how might they be overcome in the future? How do consumers of the project’s final report rate its usefulness? How helpful is the project’s final report to the court improvement process?

LESSONS LEARNED

Planning an evaluation of court performance and judicial workload takes a lot of preparation and cooperation. This *Guide* and *Toolkit* were informed by our experiences evaluating several courts’ performance and judicial workload. What follows are the lessons we learned and found were common in most sites.

Lessons Learned

PLANNING LESSONS	<ul style="list-style-type: none">• Convene Committee• Determine Scope• Determine Goals
IMPLEMENTATION LESSONS	<ul style="list-style-type: none">• Adapt Instruments• Staff Briefing• Analyze Data• Interpret Results• Examine Legal Context• Example
WRAP-UP LESSONS	<ul style="list-style-type: none">• Restate Goals• Explain Results• Future Planning

PHASE 1: PLANNING LESSONS

CONVENE THE PLANNING COMMITTEE AND BEGIN PLANNING MEETINGS

Effective communication among those individuals involved in the study is crucial to its success. Court administrators, judges, and judicial officers need to effectively communicate their interests to information technology staff, who may have to talk with child welfare agency personnel. Facilitating that kind of communication requires organizing a planning committee to guide

and oversee the entire process. Why? To ensure some degree of oversight (hence, utility), ensure that the needs of the relevant parties will be addressed, and facilitate agreement among all those involved about what they are trying to achieve. For the most effective communication, make and distribute a “contact list” either before but certainly during the meeting. Include names of all committee members, each individual’s responsibility areas, addresses/e-mails, and phone and fax numbers.

- **Think about Including the Following People as Part of Your Planning Committee**
 - **Judges and judicial officers** – they have the needed firsthand knowledge of what’s required to process abuse and neglect cases. They also can provide leadership necessary to obtain buy-in from other stakeholders and systems.
 - **Court administrators** – they hold positions that allow them a more systemic view of case processing and other court functions affecting case processing. They may facilitate access to management information systems and other data sources.
 - **Court administrative staff** – they have firsthand knowledge of court administrative processes that have direct bearing on case processing and overall court operations.
 - **Court clerks and their staff** – they have firsthand knowledge of docketing and calendaring processes, as well as knowledge of caseflow and related administrative functions.

For the most effective communication, make and distribute a “contact list” either before but certainly during the meeting.

- **Court information technology staff** – they have firsthand knowledge of what the court’s management information system holds, how its contents relate to what the court is trying to assess, and the knowledge and ability to extract such information from the system. They will also be able to assess whether the “system” has the capacity to analyze the kind of data the study requires, and whether there will be a need to share information from other systems (for example, the child welfare agency’s management information system).
 - **Agency representatives** – they have the knowledge of processes external to the court that might affect court functions and case processing. Participation of the administrative level of the child welfare agency may be necessary to facilitate information sharing across management information systems.
 - **Agency information technology staff** – they have knowledge of the child welfare agency’s information systems and what those systems are able to contribute to the project.
 - **Policy analyst or researcher** – where available, as they have the most inclusive knowledge of research methods and data analysis procedures.
- **Encourage Participation.** Engage *all* stakeholders in the need for the study and encourage them to participate to the fullest. DO explain what the purpose of the study is and what the goals are. DO NOT assume everyone will be “onboard.” Remember, people may feel their jobs or way of working will be threatened by the study.
 - **List Information Needed from Stakeholders.** Stakeholders will need to get the committee information for review before the study begins. Relevant reports (for example, annual, federal, local, progress reports and electronic data reports) need to be reviewed by all involved; court rules, administrative rules, and state statutes need to be familiar to committee members; previous relevant studies related to workload and court performance will help inform committee members; the court’s annual budget is a must for funding and financial planning issues; and the procedure for pulling hard cases and MIS data for review will need to be clarified.
 - **Review Funding Issues.** Take another look at your funding issues and at your court’s resources. Always make sure your budget can support your goals. If it can’t, one or the other will have to be revised.
 - **Organize Observations, Interviews, Case File Reviews, and MIS Review.** Review and document why a particular court, set of judicial officers, or jurisdiction has been chosen for the evaluation – document the site selection and case sample selection criteria. Consider a rotating sample of sites and case events so as not to burden a small proportion of courts unnecessarily. Determine and document what kind of data will be used by looking at the quality of data the court can provide via their MIS/electronic information system, case file reviews, interviews, and court observations. The availability of such data can be vital in determining which courts to evaluate and the nature of the sample of cases to be assessed.

- **Schedule Conference Calls and Planning Meetings.** At the end of every planning meeting schedule the next meeting. Use subsequent conference calls to develop agendas for other conference calls and further meetings. Use both conference calls and meetings before the court study begins, during its implementation, and after it's completed – this is true whether on-site staff will be doing the assessment or outside consultants will be brought in to conduct the study. Designating a court liaison for the study to be the “go to” person for issues and questions during the study makes for an easier and more expedited procedure. If outside consultants are brought in, the consultant team should be encouraged to have a single liaison for the court stakeholders to go to, as well. Require input from all stakeholders during the evaluation.
- **Transcribe Meetings and Conference Calls.** This is important not only because it enables the committee members to go back and pull out key ideas, but also because it enables the committee to reach consensus more quickly. Each transcription should include, at a minimum, the date, location, attendees, and a list of all the issues discussed or debated, along with any solutions, and next steps. Another useful tool is a “cheat sheet” for meeting participants that outlines the study’s goals and invites input from the stakeholders. Be prepared to revise this “cheat sheet” as you progress in your planning. A good “cheat sheet” will encourage discussion and participation – you might want to include such issues as cost-benefit analysis, human resources impact, and issues concerning operations and process of the study.

DETERMINING THE SCOPE OF YOUR STUDY

Determining the scope of your study means reaching agreement on how long you expect the study to take, who will be involved, what each individual’s area of expertise is, who will be responsible for each task and who will be overseeing each task, and finally, whether you’ll bring in outside people to conduct any portion of the study. This includes whether outside people will be conducting interviews, court observations, or case file reviews. If so, how will you select these individuals and who will “train” or brief them on local procedures, court protocol, and the specifics of interviewing, observing, and conducting a file review? Sketching out the process and players will begin to generate a realistic estimate of the scope of your project.

The planning committee should have people who are both very familiar with the court and represent a variety of perspectives (for example, different stakeholder groups). Committee members should have insight regarding the court’s strengths and weaknesses.

There are several ways the committee can gather input from stakeholders regarding the scope and goals of the study. First, using the core court performance measures and judicial workload assessment process outlined in earlier chapters of this *Guide* as an outline, the committee can structure the scope and goals of the study, adding areas of focus to capture information about specific jurisdictional concerns. An informal *survey* may also be developed asking other system stakeholders to indicate what they believe problem areas are, or asking them to rank order problems the court may face from a list provided by the committee. Members of the committee can also use *informal telephone calls* with appropriate

individuals. Finally, *focus groups* and *meetings* with appropriate individuals is another method of collecting input about the scope and goals of the study.

DETERMINING THE GOALS OF YOUR STUDY

In all likelihood the primary goal of your study will be to obtain some realistic measure of how your court is performing in the core performance areas, whether your judicial workload is appropriate, and what impact the workload is having on core court performance. But in order to clearly articulate goals, we found it helpful to think about the following:

SCOPE

- **Do You Want to Add to the “Core” Performance Measures by Including Measures Specific to Your Court?** What do you want to measure? What is the purpose of the evaluation? Is it to assess compliance with state laws, with ASFA, with court improvement project requirements, with the child and family services review (CFSR)? To assess the impact of a pilot program on overall court performance? To conduct targeted strategic planning for court reform?

ASSESSMENT

- **Do You Want to Assess Judicial Workloads or Caseloads or Both?** What are the differences between workload and caseload? What is the impact/expected information of evaluating one or the other or both? Do you have the resources to do both? What is the purpose of measuring workload versus caseload?

OUTPUT

- **Do You Want to Produce a Report or Something Else?** Always keep in mind tangible, practical, and helpful “products” that themselves are “goals.” Do

you want to produce a process summary to help modify and adapt the study for future evaluations? Do you want to produce a formal written report? If so, for whom? A report to the state legislature may look a little different than a report to the chief justice of the court. Do you want an analysis of your MIS electronic data – again, for what purpose and who will be reading the analysis? Do you want an analysis of interviews and observations – again, ask yourself the same questions. Finally, do you want to make recommendations? Again, consider your audience as well as what recommendations your data can realistically support.

PHASE 2: IMPLEMENTATION LESSONS

ADAPT THE INSTRUMENTS FOR YOUR COURT

The Packard *Toolkit* instruments have been designed to be user friendly and, more importantly, adaptable to the specific needs, circumstances, and culture of your court. In fact, we would strongly recommend that our instruments be reviewed well in advance of the actual data collection stage of your study and modified according to your jurisdiction’s unique criteria. Examples of adaptable-areas-per-instrument-type follow.

- **Interview Instruments** can be adapted to address your court’s unique culture, structure, and even workforce composition. Consider the following when reviewing the interview instruments for potential modifications:
 - List the types of people you will need to interview based on your need to collect additional information – are additional instruments needed?

Are there additional questions these individuals need to be asked?

- Consider distributing interview protocols and questions to an “advisory committee” to obtain feedback about modifications to reflect jurisdictional issues, structure, process, etc.
- Remember the instruments are already “pretested” for judges and court administrators, but if you elect to modify them to include other stakeholders you should also consider “testing” them with a focus group.
- **Observation Instruments** can be adapted to address your court’s hearing process. While the *Toolkit* contains hearing observation checklists for most child protection hearing types, your jurisdiction may include additional hearings not encompassed in the instruments. Generic language is also used in the Packard instruments. Changes could be made to incorporate jurisdiction-specific language, which should serve to make the court observation coding process easier.

Case File Review Instruments and Management Information Systems Data Collection Protocols can also be adapted to take into account your court’s unique structure and procedures.

Consider the following regarding case file reviews (See *Packard Court Performance and Judicial Workload Assessment Toolkit* for guidance):

- How will cases for review be selected?

- Should only “closed” or completed cases be included in the sample, thus ensuring that all case events in a dependency proceeding will be covered, or should open cases be included in order to measure the current time taken to process various case events?
- Do you look at “pre-ASFA” cases or “post-ASFA” or both?
- Do you follow cohorts of cases based on their entry into the system?
- Do you have to create a separate TPR and adoption case file data sample due to your court’s structure (e.g., different judges or courts hear these two types of cases)?

Consider the following regarding the MIS data protocols and review:

- Do the data elements outlined in previous chapters of this *Guide* make sense in the context of your court management information systems (MIS) data review?
- What are your plans to use these data?
- What are the barriers to your using these data?
- If there are barriers, what needs to be done to correct them?
- What is your court system’s *capacity* to analyze the data?
- Depending upon your capacity, will you need outside assistance?
- What are the advantages and disadvantages of doing a one-time evaluation versus ongoing evaluations? And, as part of this discussion, can the court’s MIS regularly produce the information the court needs to measure performance and assess judicial workload and, ultimately, to build capacity? If not, what can be done to make it better?

We cannot stress enough how critical it is for courts to debrief their staff and other stakeholders immediately after the interviews, observations, and case file/electronic data review.

DEBRIEF AFTER DATA COLLECTION

We cannot stress enough how critical it is for courts to debrief their staff and other stakeholders immediately after the interviews, observations, and case file/electronic data review. While the debriefings should only entail initial impressions and findings, they serve two key purposes: (1) they establish continued “buy-in” and collaboration in the entire process, and (2) they elicit prompt feedback that will further help your court evaluate performance and workloads.

ANALYZE YOUR DATA

Data analysis is essentially an examination of facts and data to provide a basis for understanding, to determine relationships, and to be used as a basis for effective decision making. Individual facts and data points are important but do not usually provide an effective basis for making systemic change or for setting priorities. Prioritizing and change depend on an understanding of *relationships*, and relationships can only be derived from an analysis of facts and data.

Using “*Triangulation*” – There are strengths and weaknesses to any single data collection strategy. Using more than one approach permits the evaluator to combine strengths and to correct some of the deficiencies of any one source of data. “*Triangulation*” denotes the process of building checks and balances into a research design through multiple data collection strategies and is aimed at increasing the accuracy of the analysis.

The Packard *Guide* and *Toolkit* uses “*Methodological Triangulation*,” which involves using multiple methods to study a single issue (for example, using interviews,

questionnaires, focus groups, court observation, and analysis of management information systems electronic data or case file reviews to study court performance and judicial workload needs). Our *Guide* and *Toolkit* also advocate the use of “*Data Triangulation*” – using a variety of data sources to study the same phenomenon (e.g., interviewing people in different positions about court performance issues).

As an analytic technique, triangulation means integrating, comparing, and contrasting information from these multiple data sources to address specific questions of interest. This results in a more comprehensive picture of reality and reinforces the validity of the conclusions drawn. And where the different data sources produce statistically different results on the same issue, triangulation can point to problems in the data collection and the need for further inquiry to resolve differences.

INTERPRET YOUR DATA

The measurement and analysis of court performance and judicial workload in dependency cases are more effective if they are tailored to the specific circumstances in a jurisdiction. This includes modifying aspects of your methodology to incorporate your court’s demographic and caseload trends, as well as interpreting results within the particular operational, cultural, and legal context of your court.

Specific techniques for calculating workload and performance measures have been briefly presented in previous chapters in this *Guide* and are further discussed in the *Packard Court Performance and Judicial Workload Assessment Toolkit*. However, incorporating data on

population and caseload trends at the measurement stage ensures that the results generated will recognize the current as well as future needs of the court. Depending on local circumstances, it may also be necessary to systematically include information on pending legislation and court rules, projected increases or decreases in personnel and other resources, and expected changes in processes and procedures within the court and related agencies in the overall analysis.

CONSIDER POPULATION AND CASELOAD ISSUES

POPULATION

Population trends are one indicator of judicial workload trends. At the broadest level, population trends are highly correlated with caseload trends; therefore, projecting population statistics assists in forecasting caseload trends and other demands on judicial resources as well as ancillary services. Population statistics and projections by age group may be especially relevant for juvenile court proceedings.

The U.S. Bureau of the Census's Web site, <http://quickfacts.census.gov> or <http://factfinder.census.gov>, provides information on population by county, including age, race/ethnicity, median income, and percent change over time. In addition, the Census Bureau, through its State Data Center (SDC) Program, makes data available locally to the public through a network of state agencies, universities, libraries, and regional and local governments. The SDC also provides training and technical assistance in accessing and using Census Bureau data for research, planning and administration, and decision making. The SDC for each state can be found at <http://www.census.gov/sdc/www/>.

CASELOAD

Determination of case-filing rates by category provides for analysis of filing rate trends independent of population growth. Valid and reliable information on caseload is essential to the measurement of workload and establishing standards. Statistics on historical trends in case filings should be available from automated or manual data collections at the state administration or local court level.

While standard frameworks for projecting population and court filings are applicable to most circumstances, unique circumstances may be encountered in select jurisdictions. Typically, these may involve the lack of data or reliable historical trends upon which to base projections. In these circumstances, the collaborative judgment of the researcher and the court may be required to define an approach that is appropriate.

INTERPRETING PERFORMANCE AND WORKLOAD RESULTS

Understanding the wider context in which the dependency court functions is essential to interpreting the results of the measurement of workload and performance. By their nature, many contextual factors are not subject to the control of the court, and their impact, though real, can be difficult to measure and manipulate quantitatively. Potentially significant factors will likely vary across courts and over time, and not every factor will be equally salient for all courts. Incorporating contextual factors into the interpretation of workload and performance measurement is further complicated by the fact that linkages may be indirect or mediated by other factors or forces.

Workload and performance may be affected by the way in which the court is organized around its work and deploys its resources.

Information on contextual factors will likely be gathered from multiple sources. As noted, population data are available from the Census Bureau and its State Data Centers. Local county planning departments or regional planning agencies may also be able to provide population data and projections. Historical caseload data should be available from state court administrative offices or the local court. Information on court structure, court and agency personnel and other resources, and case-processing procedures may be available from court or agency annual reports or other system documentation. It is likely, however, that much of the detailed information required for the assessment will be collected through the interview and court observation process conducted during the assessment.

CONSIDER DEMOGRAPHIC FACTORS

Clearly, dramatic increases in population may fuel caseload growth that hampers a court's ability to process cases in a timely manner. It may also strain the capacity of the social service network to provide the services needed by families in a timely way. On the other hand, a rapidly declining population that reflects erosion of a community's economic base may also indicate a decline in the resources available to the court. The impact of population growth trends on a court's performance on the various measures related to achieving permanency in dependency cases, and timeliness in case processing, should be considered in the analysis.

Characteristics of the population also have implications for performance. Information on the socioeconomic and racial/ethnic characteristics of a jurisdiction's population may speak to the need for specialized services and

programs that will ensure that the court is accessible and responsive to the population it serves and does so in culturally competent ways. A lack of culturally appropriate materials and programs for families involved in dependency proceedings can contribute to delay in case processing and can undermine, if not negate, the effectiveness of any services being offered. Similarly, if a court is serving a linguistically diverse population, it must ensure the availability of a sufficient number of interpreters, multilingual attorneys, and other staff to ensure due process for all participants. Whether the court is equipped to respond to the social, economic, ethnic, and cultural make-up of the population that it serves should be considered in the analysis. The characteristics of the population may affect a court's performance on measures related to the timeliness of case processing.

CONSIDER THE IMPACT OF COURT STRUCTURE

Workload and performance may be affected by the way in which the court is organized around its work and deploys its resources. For example, an assumption underlying the establishment of a specialized family court is that judges will have more expertise in that area and thus make considered decisions in less time. Judges who must rotate caseloads require more time to acquaint themselves with family law. A court must have a sufficient number of judicial officers to effectively handle the dependency caseload if it is to achieve the core outcome goals of safety, permanency, and well-being. The degree of specialization, length of assignments, arrangement of calendars, use of extra-judicial officers, and availability of judicial support staff are important considerations in both shaping and interpreting the results of the assessment.

These structural factors influence the amount of overall time available for dependency cases and the time that can be scheduled for specific hearings. In turn, this may affect the overall timeliness of case processing, the time between specific case events, and the ability of the judicial officer to address the issues and make the findings mandated at each stage of the case. For instance, a judge who has access to a law clerk and other support staff to assist with the review of cases and the gathering of information prior to hearings may be able to spend less time on each hearing. In addition, these factors impact the capacity of the court to achieve a long-term perspective on the issues and needs of the family and render the consistent and informed decisions that enhance the prospects for achieving permanency at an earlier stage of the case.⁶⁸

CONSIDER CASELOAD VOLUME AND TRENDS.

Caseload volume, composition, and characteristics, viewed over time, show current demands on resources and services. They can also help to predict future demands. Changes in these factors have implications for the court’s ability to achieve permanency within recommended time lines, assure due process, and comply with overall and interim case-processing time standards. Caseload volume and complexity determine in part how a court organizes and deploys its resources around particular case types. Sudden influxes of new cases or a rapidly growing number of pending cases may signal the need to add new resources or reallocate existing resources on at least a temporary basis.

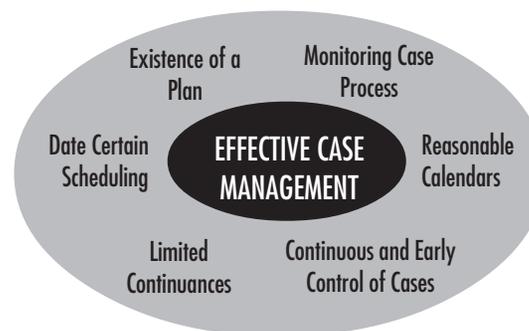
CONSIDER CASEFLOW MANAGEMENT

Timeliness is a core process goal in dependency cases and a number of key measures for this goal have already been

outlined in previous chapters of this *Guide*. Adhering to fundamental principles of caseflow management is key to achieving compliance with the time standards outlined in those measures.⁶⁹ Caseflow management is also essential in dependency cases if permanency planning efforts are to be successful. Delay and uncertainty undermine the responsiveness and accountability that are goals of the system. Establishing the firm expectation that events will occur according to established schedules and time lines is especially important when cases involve multiple players.

Consideration of caseflow management procedures ensures that your assessment process differentiates between workload and performance issues that are due to resources and those that may be the result of less than efficient case management processes. Courts that have already incorporated the basic tenets and techniques of caseflow management into their operations and created a system-wide culture that supports judicial control of case progress are presumably better situated to respond to the rigorous time frames and other demands of AFSA.

Among the key components of effective case management systems are:



⁶⁸ See the *Packard Toolkit* for a judicial worksheet instrument containing questions about the structure for handling dependency cases and the effectiveness of the calendaring system.

⁶⁹ The “Case Processing Time Standards in State Courts, 2002-03” are now available on the National Center for State Courts Web site at http://www.ncsconline.org/WC/Publications/KIS_CasManCPTSPub.pdf.

⁷⁰ P.L. 105-89, amending 42 U.S.C § 601 et seq.

⁷¹ *Supra*, note 5.

⁷² Adoption and Safe Families Act of 1997 (P.L. 105-89) signed into law November 19, 1997, amending Titles IV-B and IV-E of the Social Security Act.

⁷³ U.S. General Accounting Office, *Juvenile Courts: Reforms Aim to Better Serve Maltreated Children* (Washington, DC: U.S. General Accounting Office, 1999), p. 8.

⁷⁴ U.S. Department of Health and Human Services, Children's Bureau, *Guidelines for Public Policy and State Legislation Governing Permanence for Children* (Washington, DC: U.S. Department of Health and Human Services, 1999), pp. 1-5 – 1-6.

Innovative approaches are needed to achieve the goals of safety, permanency, and well-being.

The instruments contained within the *Packard Court Performance Measurement and Judicial Workload Assessment Toolkit* include survey forms (i.e., interviews, questionnaires, and focus group guides) developed for judges, attorneys, and other personnel, which incorporate a number of questions related to caseload and calendar management.

Understanding the wider demographic, cultural, and organizational context in which the court functions is essential for interpreting results of your performance and workload measurement. And it is critical in designing and managing improvement strategies. Contextual factors may have both short- and long-term, as well as direct and indirect, consequences for the court system. In addition, elements of the court's structure, process, and resources, and those of related agencies and offices, may affect operations and the delivery of services.

EXAMINE THE LEGAL CONTEXT

This section discusses the legal framework guiding the operations of child welfare courts and its impact upon court performance and judicial workload. An analysis of the legal context serves two purposes with respect to data analysis and interpretation. First, it helps to provide a thorough understanding of the state's statutes and the related legal requirements, time lines, and significant court milestones in dependency cases. Second, the legal context analysis highlights a discussion of the state statutes and court rules in comparison to the mandatory provisions outlined in (1) the federal legislation of ASFA⁷⁰ and (2) the aspirational provisions envisioned in the *Resource Guidelines* and *Adoption and Permanency*

Guidelines.⁷¹ This will enable a court to assess how the laws of its state fare in relation to nationwide requirements and nationally recognized "best-practice" standards.

THE ADOPTION AND SAFE FAMILIES ACT OF 1997

Of primary concern is the impact of the Adoption and Safe Families Act.⁷² ASFA was, in part, "a response to the fact that more children were entering the foster care system than were exiting."⁷³ This landmark legislation clearly and unequivocally established the national goals of safety, permanency, and well-being for children in foster care. Five principles underlie ASFA, evolving from some of the assumptions underlying Adoption 2002, and these apply to professionals working with families through public and private agencies, as well as state courts. These principles are *safety is the paramount concern that must guide all child welfare services; foster care is temporary; permanency planning efforts should begin as soon as the child enters care; and the child welfare system must focus on results and accountability*.

Innovative approaches are needed to achieve the goals of safety, permanency, and well-being.⁷⁴

ASFA necessitates timelier, decisive, and substantive hearings, and more frequent court and administrative case reviews. It also requires a focus on outcomes and performance reports, and stresses both court and child welfare system accountability. ASFA also stresses the need for collaboration and community partnerships that are focused on child safety and timely permanency. The implementation of ASFA presents a number of major

challenges for state courts and the child welfare system. Because state compliance with the law is a condition of state eligibility for funding to public child welfare agencies, ASFA places new demands on state court resources. Moreover, the passage of ASFA also significantly increases the role of the judiciary, as well as the agencies and advocates, throughout the processing of the case and ultimately places responsibility for compliance and good outcomes for children and families squarely on the shoulders of the court.

THE NCJFCJ RESOURCE GUIDELINES

The National Council of Juvenile and Family Court Judges (NCJFCJ) has published *Resource Guidelines* that identify a series of “best practices” courts should follow in processing child welfare cases. The *Resource Guidelines* identify the (1) purpose and intent of key court hearings, (2) the timing of key court hearings, (3) the minimum duration of key court events, and (4) the key decisions

NCJFCJ Resource Guidelines Suggested Minimum Time Allocated	
Type of Hearing	Minimum Time
Preliminary Protective Custody	60 min
Adjudication	30 min
Disposition	30 min
Review	30 min
Permanency	60 min
Termination of Parental Rights	60 min
Adoption	30 min
Minimum Total Time/Case	300 min

the court should make during each hearing.⁷⁵ Some consideration should be given to how a state’s provisions compare to the processes envisioned in the *Resource Guidelines* when considering “aspirational” or benchmark goals for court performance measurement.

⁷⁵ *Supra*, note 5.

STATE AND LOCAL STATUTES OR COURT RULES

ASFA requires that, as a condition of federal funding, courts must process child welfare cases in accordance with specific minimum criteria. Each state, however, can promulgate laws that are more demanding than ASFA so long as the restrictions do not violate the constitutional rights of the parties. The *Resource Guidelines*, while aspirational, represent the basic requirements for optimal case processing, judicial activity, and oversight. The review and comparison of state statutes and court rules to ASFA and the *Resource Guidelines* is helpful to determine how a court compares to the processes envisioned in the *Resource Guidelines* and the mandatory provisions of ASFA.

EXAMPLE: THE IMPACT OF RESOURCE GUIDELINES CONSIDERATIONS ON WORKLOAD

The breadth of the activity suggested by the *Resource Guidelines* during each hearing requires that the court dedicate a significant period of time for the suggested activity for each hearing type. The following table indicates the suggested minimum times for each case type and the total time (300 minutes) for each child welfare case that the court processes. The *Resource Guidelines* do not distinguish between contested and uncontested hearings and involve the minimum suggested time to be allocated for a hearing event.

INTEGRATED

Integrated with other information derived during the evaluation process, such as the interview and court observation results, in a process of methodological and data “triangulation.”

ANALYZED

Analyzed and interpreted in the context of the court, ensuring that findings are relevant and useful to your jurisdiction.

COMMUNICATED

Communicated in clear, tailored, and useful ways to the study’s various audiences.

INCORPORATED

Included in plans and strategies for systems change.

In most jurisdictions, however, the volume of cases and limited judicial resources precludes the allocation of time recommended by the *Resource Guidelines*. Significant time allocated to earlier hearing events may, however, reduce the judicial time needed for later hearings because more issues have been identified and resolved early on in the life of the case. For example, in several jurisdictions preliminary protective hearings are handled very quickly, often taking merely between 10 and 20 minutes. Procedures could be developed, however, for longer and more meaningful preliminary protective hearings. While a one-hour preliminary protective hearing might be too time-consuming to be held as a routine event in every case, the tactical use (especially in difficult cases that would otherwise be inordinately time-consuming for the court at a later stage) would be very productive.

A review of the legal context in which the court operates can help a court to identify the strengths and weaknesses of its case-processing framework and the resulting impact on court performance and judicial workload. Once the court has identified these strengths and weaknesses, the court can make modifications to its operations that will more closely align the court with optimal case processing and improved outcomes for children and families.

PHASE 3: WRAP-UP LESSONS

In the final phase of the project all your hard work gets memorialized. It is vital to your study and the overall evaluation process that your findings be integrated, analyzed, communicated, and incorporated effectively. The foundation for successfully completing your evaluation is actually established in the first two phases of

the evaluation – planning and implementation. By clearly identifying the goals of the study, involving all of the key stakeholders, achieving consensus on the questions to be asked and relevant measures, and adapting the instruments and protocols to fit the court environment, you will be able to produce valid, useful, and convincing results. Equally important to consider during this phase, but again, to have been considered from the very beginning, are the steps you will need to take to generate a report, an exploration of the potential uses for the data generated, and the best way to publicize and present the data.

The following are suggested areas of discussion you should think about addressing in your report or briefing (See *Packard Court Performance and Judicial Workload Assessment Toolkit* for more guidance.) These were gleaned from our experiences evaluating a number of courts as a component of the Packard Project. And while these topics are correctly listed under “Phase 3” of the overall project, remember that being aware of their importance from the beginning of the planning process through implementation and into completion will lead to better informed results.

WHAT ARE YOU TRYING TO EXPLAIN BY DOING THIS EVALUATION?

PERFORMANCE MEASUREMENT

- What have you learned about your court’s performance?
- What performance measures do you lack, and which ones will you need in the future?

- What do you need to do for future performance measurement? Explore options for routinely obtaining the measures from existing automated information systems. Develop easily accessible, routine reports that allow “at-a-glance” assessment of current operations and changes in performance over time.
- Judicial Workload Assessment needs?
- What do you need to do for future workload analysis?

WHAT’S THE BEST WAY TO EXPLAIN YOUR DATA?

Consider who your audience is: this will inform how you can effectively explain what the data means. Think about the different avenues for communicating the data (reports, articles, symposia, debriefings, and the like). Include the importance of acknowledging the need to document the study and the methods used. Include the importance of acknowledging the weaknesses and limitations of data sources used and the data collection procedures implemented.

WHAT’S THE BEST WAY TO ADDRESS FUTURE PLANNING?

Of particular concern to courts should be the need for action planning – those specific actions that respond to short- and long-term objectives. Action plans include details of tasks to be accomplished, resource commitments required (people, time, and money), assignments, and time lines for completion. Action plan development represents a critical stage in planning when strategic objectives and goals are made specific so that effective court-wide understanding and deployment of resources are possible. (See *Packard Court Performance and Judicial Workload Assessment Toolkit* for sample action-planning worksheets.) Consider planning future

Consider planning future evaluations around pervasive problem areas, such as perceived lack of resources and lack of cooperation between and among systems.

evaluations around pervasive problem areas, such as perceived lack of resources and lack of cooperation between and among systems.

The court is only one component of a large and complex system that is assembled to respond to abuse and neglect cases. An adequately resourced social service system that can deliver services immediately to families in crisis and trained and committed guardians *ad litem* or attorneys for children, indigent parents, and the social service agency are also important to achieving successful outcomes in dependency cases.

The availability of ancillary specialized services and programs and the effective use of volunteers and other no-cost resources are critical to developing a comprehensive response to the problem of abuse and neglect. Absence of needed services or failure to make appropriate matches of clients to services can undermine timeliness and the achievement of permanency. Collaboration between the courts and service providers not only provides the opportunity for more comprehensive treatment for families, but also provides the court with an entry point into the community. The court learns what services are available and what referrals are appropriate. The service providers learn what the court expects and what type of reports and other information are helpful to the court. Courts may find it useful to conduct a periodic “service inventory” by taking the time to review the services available in the area and to make this information available to court staff and families.

Future evaluations may include more comprehensive studies that identify an inventory of services available to children and families involved in dependency cases and a determination if the absence of any particular resource is causing delay in case processing or hampering resolution of certain cases. Among the services of interest are supervised visitation; alcohol and drug assessment and treatment; domestic violence intervention; mental health services; medical services; housing; parenting skills education; and intensive home-based services. Ideally a comprehensive evaluation of performance would also consider if the court has maximized opportunities to access the broader resources of the community and if ongoing partnerships have been formed.

