A New Method of Assessing Judicial Workload in Juvenile Dependency Cases

By Alicia Summers, Stephanie O. Macgill, Sophia I. Gatowski, Jesse R. Russell, and Steve Wood

ABSTRACT

This article describes a new method for calculating judicial workload in dependency or child abuse and neglect cases. In contrast to traditional judicial workload methods, the method described herein produces estimates of judicial workload that take into account the complex role of the juvenile dependency court judge—a role that includes both on- and off-the-bench activities. The method provides workload estimates that give guidance to courts not only about the minimally sufficient judicial resources needed to accommodate current caseload needs, but also what level of judicial resources would be required to hold substantive dependency court hearings that comport with nationally recognized practice recommendations. The article reviews commonly used judicial workload methods, outlines the new method, and uses a pilot of the method as an example of how the method works in practice. Broader implications of this workload method are also discussed.

Identifying factors that might contribute to delays in timely permanency for abuse and neglect victims has been a focus of many juvenile dependency researchers (e.g. Benedict & White, 1991; Courtney, 1994; Fernandez & Lee, 2011; Harris & Courtney, 2003). This focus comes from not only an effort to identify ways to improve outcomes for children and families, but also an interest in helping juvenile dependency courts meet their federal requirements and state goals in this area. The Adoption and
Safe Families Act of 1997 (ASFA) requires that children removed from their homes due to allegations of abuse or neglect be placed in safe, permanent homes within a stricter timeframe than prior federal laws, hoping to decrease the time children spend in non-permanent homes. Further, the U.S. Department of Health and Human Services analyzes states’ performance on seven outcome measures related to safety, permanency, and well-being in the Child and Family Services Review (CFSR), and imposes financial penalties on states whose outcomes fall below federal standards. Overall, the federal government has not been satisfied with the CFSR performance of most states (Meckler, 2003). The most recent CFSRs, conducted between 2007 and 2010, found that no state achieved substantial compliance for six of the seven measures, although 10 states achieved substantial compliance for the well-being outcome related to children achieving appropriate services to meet their educational needs (U.S. Department of Health and Human Services, 2011). States’ shortcomings with respect to federal mandates have led researchers to attempt to identify factors that may facilitate or inhibit the achievement of safer and timelier permanent placements. One emergent factor in the literature has been judicial workload.

Judicial workload has gained attention because increases in caseloads, without accompanying increases in judicial personnel, can compound the burden on courts (Outley, 2006). This burden is particularly troublesome in dependency cases, which are often complex and require multiple hearings across the life of the case (Hardin, 1996). Early studies on judicial workload found that between 52% and 79% of juvenile dependency judges cite time constraints and overcrowded court dockets as the biggest challenges to successfully completing their judicial duties both on and off the bench (Dobbin & Gatowski, 2001; Fostering Results, 2004). In some cases, these challenges can lead to delays in achieving permanency (Fostering Results, 2004). Because of these findings, states like Minnesota (National Center for State Courts, 2003), West Virginia (National Center for State Courts, 2006), Texas (National Center for State Courts, 2007), and Michigan (National Center for State Courts, 2011) began implementing workload studies to examine the number of judicial officers necessary to handle juvenile dependency caseloads thoroughly and efficiently.

Workload assessments typically utilize a weighted caseload method for determining judicial workload and judicial resource needs. This method focuses on how much time a judge has to hear the case and how much time is required to oversee specific types of hearings. This approach to judicial workload assessment, however, is insufficient for two reasons. First, it does not fully account for the complexity of juvenile dependency work and its demands on judicial officers both on and off the bench. Second, it does not sufficiently account for differences in hearing quality; it evaluates only judicial work from a perspective of what is required to be minimally sufficient, not the work that is required for judges to adhere to nationally recognized practice standards for conducting thorough dependency court hearings.

Herein, we describe a new method for judicial workload assessment, which not only provides a more accurate estimate of judicial resource needs, but also accounts for the time needed to comport with best practice standards in dependency court hearings. After discussing juvenile dependency workload and previously used judicial workload
assessment methods, we outline the data elements and instruments used in the new method. We then discuss implementation of the method and utilize a pilot workload assessment to demonstrate how the method works in practice.

**JUVENILE DEPENDENCY WORKLOAD**

Juvenile dependency cases are complex and unique. They require more social services, collaboration between courts and child welfare agencies, and community involvement than most other types of cases (Fiernone & Salyers, 2005; Miller, 1999). Juvenile dependency cases require active and consistent court oversight, multiple hearings and frequent court reviews, and a broad and active scope of inquiry from the bench, while staying within demanding state and federal timeframes (Hardin, 1996). Indeed, “[t]he role of the juvenile court judge combines judicial, administrative, collaborative, and advocacy components” (Edwards, 1992, p. 25).

The complex demands of dependency cases, the high volumes of caseloads, and the state and federal mandates require adequate judicial resources to ensure that dependency courts can provide safe and timely permanency for children and families. Thorough hearings may be related to better placement outcomes for children and youth (National Council of Juvenile and Family Court Judges [NCJFCJ], 2011a). In one study, judges who held hearings that are more thorough also had more family placements and fewer non-relative foster care placements (NCJFCJ, 2011a).

Substantive discussion of key dependency issues, high levels of stakeholder engagement, and consistent and active judicial inquiry distinguish thorough hearings. Key issues for discussion include, but are not limited to, parties who should be present, the applicability of the Indian Child Welfare Act (ICWA), placement of the child, reasonable efforts made to prevent removal or return the child home, and the appropriateness of services offered (NCJFCJ, 2011a). Best practice guidelines recommend that judicial officers actively ask questions of social workers, attorneys, parents, and children (if present) to fully discuss each issue, hold parties accountable, and encourage meaningful engagement in the dependency process (NCJFCJ, 1995). Despite these practice recommendations, previous studies have found that many dependency judges cite time constraints as their biggest challenge (Dobbin & Gatowski, 2001; Fostering Results, 2004). Accordingly, overburdened judicial officers may not have adequate time to conduct thorough hearings and fully engage all parties through meaningful discussion.

**JUDICIAL WORKLOAD METHODS**

Three traditional methods measure judicial workload: the weighted caseload method, the Delphi method, and the normative method (Dobbin & Gatowski, 2001; Hurst, 1999). Of the three, the weighted caseload method provides the most in-depth and accurate assessment of workload. Weighted caseload estimates involve
identifying case events, determining the frequency of each event, and summing the
time that a judicial officer spends on the event. This formula yields a weighted case
value. This approach captures a measure of case complexity in relation to judicial
workload estimates, but it does not take into account a standard of practice. Further,
this method does not take into account the time that judges may need to spend off the
bench preparing for or following up on hearings, potentially resulting in substantial
underestimates.

The Delphi method of workload assessment involves gathering expert opinions
regarding the average amount of time spent on each case event, from which case weights
can be developed. These case weights are used with judicial hour estimates to predict
judicial resource needs. Drawbacks for this approach are the same as the weighted
caseload approach and include an additional limitation: the measure can be subjective and
overly dependent on the particular experience of the focus group if used without adjunc-
tive methods.

The third approach is the normative method—a cost-efficient and simple method
of calculating judicial workload. The normative method compares similar jurisdic-
tions, dividing the number of cases by available judicial resources in each jurisdiction.
This formula produces an estimate that identifies the number of judicial officers per a
set number of cases. The primary weakness of this approach is the assumption that
average practice is appropriate practice. Like the weighted caseload approach, the
approach does not account for whether current practice comports with recommended
practice.

These three methods can be useful for establishing baseline data and helping a
jurisdiction begin to quantify how caseloads relate to workload. However, they provide
an incomplete understanding of workload. They fail to consider critical judicial activities
occurring off the bench and thus may substantially underestimate the number of judges
needed to handle current caseload demands effectively and efficiently (Dobbin &
Gatowski, 2001). In dependency cases, judges may spend quite a bit of time off the bench
reviewing reports for upcoming hearings, as well as other administrative, training, or
meeting tasks related to their dependency work. Further, these methods account only for
hearing time, and not for hearing quality. They overlook the fact that actual hearing time
may not be sufficient to address the complex range of case issues relevant to dependency
cases. At issue is content validity, which evaluates the efficacy of the measure at tapping
into the concepts as they are currently defined (de Vaus, 2001). For example, an arith-
metic exam that includes only subtraction problems would not be a valid measure of
overall arithmetic skill. Similarly, workload methods that do not consider quality hearing
and off-the-bench factors when estimating a jurisdiction’s needs do not capture all aspects
of judicial workload in juvenile dependency cases.

To address these deficiencies, the National Council of Juvenile and Family Court
Judges designed a new method of calculating dependency workload. This method builds
and expands upon prior workload assessment methodologies by accounting for the range
of activities judicial officers may dedicate to juvenile dependency off the bench, while also
accounting for the quality of practice by examining the depth and breadth of dependency
hearings.
WORKLOAD MEASURES

This new judicial workload method entails multiple data elements. Each element is discussed in detail below.

Data Elements

The workload method requires five specific data elements: (1) time spent on dependency activities (both on and off the bench); (2) average hearing length relative to hearing quality; (3) number and type of dependency hearings per year; (4) number of full-time equivalent judges per jurisdiction; and (5) hours of a typical judicial work year.

The first data element—time spent on dependency activities—captures not only time that judges spend overseeing a court hearing, but also time spent off the bench preparing for and following up from a hearing. Due to the complex nature of dependency cases, status changes and reports are typically submitted to the court prior to hearings. Time spent reviewing these reports cannot be captured with an on-the-bench measure. Therefore, this data element includes three elements of judicial time—time spent on the bench overseeing hearings, time spent preparing for dependency hearings, and time spent following up on dependency hearings. Other off-the-bench activities, such as time judicial officers spent doing administrative activities, receiving training, or conducting community outreach, are also important and provide additional context to the assessment. These activities are often required of judges and take time away from the bench. As such, they are helpful in determining how many hours and days judicial officers actually spend conducting hearings.

The second data element is average hearing length, in minutes. This data element is calculated relative to hearing discussion. That is, we calculate how long a hearing with thorough discussion lasts (in minutes), as well as how long a hearing lasts when there is limited discussion. Data element two requires hearing observation data that can be gathered by judges or other observers. By noting hearing start and end times, and by coding the level of discussion during hearings, this measure includes an estimate of hearing length as well as a measure of the thoroughness of the hearing.

The third data element is the number and type of hearings (e.g., preliminary protective, adjudication, review, permanency) per year that juvenile dependency judges oversee. Different types of hearings may require more judicial time and some may occur more frequently than others, depending on state laws and local practice. These numbers are often available through administrative information systems within the court, or can be estimated based on the number of hearings in a typical month. To control for case volume fluctuations from year to year, courts can average the number of case events across three or four years, if possible.

The fourth data element of interest is the full-time equivalent (FTE) judges, or current judicial resources. The amount of time each judicial officer spends overseeing dependency cases contributes to the total FTE judicial resources in a jurisdiction. For example, if a jurisdiction employs six full-time judicial officers, each of whom spends 20% of his or her time on dependency cases (one full day in a five-day work week), that
jurisdiction has a total of 1.20 FTE judicial officers. Courts can easily calculate the judicial officer FTE level by first determining the total number of judges currently overseeing dependency cases and then adding the percentage of each judges’ time devoted to hearing dependency cases.

The fifth data element is the number of hours a judicial officer is available to work in a given year. Determining this number can be approached several ways. Courts can assess the actual number of hours each judicial officer worked in a year, or courts can use an estimate of hours in a typical work year. A review of prior workload assessments found that a typical judicial year consisted of 215 days, or 1,720 hours (Kleiman & Lee, 2011; National Center for State Courts, 2003). Another approach is to use an estimate of 2,000 annual working hours, which reflects the assumption that judicial officers work 50 weeks a year, 40 hours a week (NCJFCJ, 2011c).

### Instruments

To determine time spent on dependency activities (the first data element), an off-the-bench log is necessary (see Figure 1). This instrument is used to document the judges’ various activities (e.g., hearing preparation and follow-up) and the allocation of time for each activity. Judges complete this log every day of the study period, regardless of the dependency docket.

Determining average hearing length and average hearing quality (data element two) requires an on-the-bench log (see Figure 2). This instrument is for judges, or independent observers, to complete for each hearing they conduct during the study period. The instrument collects data on hearing length and the level of discussion of key items, as well as whether the hearing was contested or continued, the parties present, and if there was enough time for discussion. The key discussion items were selected from the Resource Guidelines: Improving Court Practice in Child Abuse and Neglect Cases’ recommendations for best practices in dependency court hearings (NCJFCJ, 1995). The instrument asks judges or observers to identify the level of discussion of each issue on a scale of 1 to 4, with 1 = no discussion, 2 = statement only, 3 = sufficient discussion, and 4 = substantive discussion (i.e., thoroughly discussing the issue). Sufficient discussion constitutes more than a statement but is not as in-depth as possible (e.g., other possible subtopics

<table>
<thead>
<tr>
<th>Activity</th>
<th>Time (in minutes)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Court Hearings (includes waiting time, consulting with attorneys and all dependency hearings)</td>
<td></td>
</tr>
<tr>
<td>Off-the-Bench Hearing Preparation (e.g., review of relevant materials, planning, preparing orders)</td>
<td></td>
</tr>
<tr>
<td>Off-the-Bench Hearing Follow-Up (e.g., preparing orders, case-related meetings, reviewing materials)</td>
<td></td>
</tr>
</tbody>
</table>

FIGURE 1. Excerpt from the off-the-bench time log instrument
could emerge or more information could be helpful to the discussion), and substantive discussion exhausts or nearly exhausts the topic. Hearings with a sufficient level of discussion have a discussion average between 2.75 and 3.25 (this allows researchers to capture variation around the mean of 3.0), and substantive hearings have average discussion levels between 3.26 and 4.0.

**IMPLEMENTING THE METHOD**

Implementing the judicial workload method requires consideration of several factors, such as judicial buy-in and cooperation and study period length. Data elements one and two (time spent on dependency activities and average hearing length) require a significant commitment of all judicial officers in a given site for reliable and valid data collection. Judicial officers complete the off-the-bench time log every day and complete the on-the-bench time log for every dependency hearing. Completion of the logs is a critical task, and without judicial commitment the workload method cannot be properly implemented. To prepare for completion of the judicial time logs, coordinated training on the instrument and on coding for all participating judges prior to use is advisable. Also prior to use, a systematic consideration of coder reliability and any potential coding discrepancies is advisable.

Another important element of implementing the judicial workload method is determining the time period for data collection, or how long judges may need to
complete the on- and off-the-bench logs. This determination can vary depending on jurisdiction size and caseload but must be sufficient to obtain a representative sample of cases. A two- to four-week data collection period in jurisdictions that oversee dependency dockets every day, and a minimum four-week period in jurisdictions that do not, may be a reasonable amount of time.

CALCULATING JUDICIAL WORKLOAD

To help explicate the method, the NCJFCJ pilot workload project from Washington State provides an example. With support from the Washington Administrative Office of the Courts, the assessment began in three jurisdictions (where data elements one and two were calculated) and then was expanded statewide. The results presented below were informed by the pilot, and full presentations of the method development, study timeframe, and results are available in several reports (Dobbin, Gatowski, Russell, & Summers, 2010; Dobbin, Gatowski, & Summers, 2010).

The first vital piece of information collected was judicial time spent on dependency activities (data element one). The off-the-bench logs indicated that judges’ time spent on dependency cases was nearly equally spent in hearings (57% of the time) and on preparing for (39%) or following up from hearings (4%). That outcome means that for every hour spent in a hearing, the judge needed an additional 45 minutes off-the-bench to prepare and follow up.

The average hearing time relative to hearing quality (data element two) was the second piece of information calculated. In the three pilot sites, average discussion level was 2.7 (or nearly sufficient). The hearing time needed for sufficient discussion was then estimated by selecting only the hearings that had a sufficient discussion level (between 2.75 to 3.25) and calculating the average hearing length. The same procedure was used to determine hearing lengths when discussion was substantive (between 3.26 and 4.0) (see Table 1).

<table>
<thead>
<tr>
<th>Hearing Type</th>
<th>Average Hearing Length for Sufficient Discussion (in minutes)</th>
<th>Average Hearing Length for Substantive Discussion (in minutes)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preliminary Protective</td>
<td>20</td>
<td>63</td>
</tr>
<tr>
<td>Adjudication/Disposition</td>
<td>11</td>
<td>30</td>
</tr>
<tr>
<td>Review</td>
<td>16</td>
<td>24</td>
</tr>
<tr>
<td>Permanency Planning</td>
<td>20</td>
<td>32</td>
</tr>
<tr>
<td>Motion</td>
<td>20</td>
<td>20</td>
</tr>
</tbody>
</table>

TABLE 1
Average Length of Hearing by Discussion Level
Data element three (number of hearings by type) was determined through administrative systems in each jurisdiction. Taking the average number of hearings per year and per type, multiplied by the estimated hearing length for each hearing type, provided an estimate of the total number of hours spent in dependency hearings annually. The additional time needed to prepare for and follow up on hearings was then added to the total number of annual hearing hours. This number provided the total hours that a judge would need to devote to dependency cases. The total number of dependency hours was then divided by 2,000 (data element five). The result was the FTE need, from which the actual FTE (data element four) was subtracted to determine current staffing proficiencies or deficiencies. Table 2 illustrates a hypothetical example to demonstrate the method of calculation.

Table 2 demonstrates how the data elements were used in piloting the judicial workload method. In this example, a jurisdiction would need 0.50 FTE judicial officers to meet sufficient practice workload demands. A jurisdiction with a current FTE of 0.30, and an estimated need of 0.50 FTE, would have a current FTE deficit of 0.20. In other words, the jurisdiction would need an additional judge overseeing dependency (or increase the time of the current judge) for an additional 8 hours each week (20% of a 40-hour work week) to meet sufficient practice needs.

The same estimate was calculated with substantive discussion estimates. Hearing length for substantive discussion nearly doubled for each hearing type, resulting in an
estimated FTE of 0.95 in the same jurisdiction. That is, taking into account the quality of the hearing, the jurisdiction would need 0.95 FTE judicial officers to conduct thorough, best practice hearings. FTE need score could then be compared to the current FTE to determine what additional level of judicial resources would be needed to conduct substantive hearings.

**DISCUSSION**

This paper describes a new method of assessing workload in juvenile dependency cases, one that acknowledges that workload needs vary considerably between sufficient practice and substantive, high-quality practice. Unlike prior workload estimates, this new method accounts for the complexity of juvenile dependency practice by adding dimensions of time needed off the bench to prepare for and follow up on hearings, and by accounting for the hearing's quality. Research has demonstrated the importance of thorough hearings; when hearings are thorough, children may have better placement outcomes (NCJFCJ, 2011a). Substantive hearings also allow for better engagement of parties, as there is more discussion of case issues and more opportunity for the parents to be involved in the process. Engagement, too, has been linked to positive outcomes for families, with research suggesting that participation of parents at hearings is related to timely reunification (Wood & Russell, 2011). Including hearing quality proxies as part of the workload assessment method improves upon other methods by accounting for the case complexity and resource needs, both on and off the bench that are required to hold substantive, high-quality hearings that may help to expedite the process and achieve timelier and safer permanency for children.

Research has also linked judicial workload to better permanency outcomes for children. An examination of workload estimates in one state compared under-staffed, staffed, and over-staffed jurisdictions on percentage of cases to achieve permanency within 15 months of entry into care (NCJFCJ, 2011b). Jurisdictions with over-staffed judicial resources also had the highest percentages of children achieving permanency within 15 months, while the under-staffed jurisdictions had the lowest achievement of permanency.

The method presented here provides a secondary benefit as well. The information gathered on hearing length relative to quality can provide important baseline information to the jurisdiction regarding current best practices and the perception of judges about whether there was sufficient time to discuss all issues. For example, in the Washington State pilot assessment, the results stimulated discussion of means to increase resources and to identify practices and procedures that might be changed to increase efficiency of case progression. As a result, one court implemented a mediation program to help expedite the hearing process and reduce the overall number of hearings for judges and other stakeholders.

This type of workload assessment may also be beneficial in future research directions related to workload stress and burnout. As this method provides a more complex understanding of workload, including time off the bench, it may be important in
understanding the level of stress and work that dependency court judges face. Prior research has demonstrated that the demand placed on judges is related to perceptions of burnout (Tsai & Chan, 2010). Dependency court judges may be particularly susceptible to burnout when their workloads are high due to the nature of the cases they oversee. A related study, examining immigration judges who are also faced with stories of human suffering, found that higher stress and burnout were related to increased psychological and health concerns (Lustig et al., 2008). Dependency court judges, because of the nature of the cases they oversee, may be more susceptible to burnout and the negative emotional, health, and psychological consequences that stress can create. A more complete understanding of workload can help provide estimates of the number of judges needed to adequately oversee a caseload, and can be used in further research related to workload, stress, and burnout among dependency court judges.

Limitations

As noted in prior workload reports, 1,720 hours is a median estimate for a judicial work year (Kleiman & Lee, 2011). Workload estimates based on a 1,720-hour work year will be different from estimates based on a 2,000-hour work year. Identifying the state- or court-specific judicial work year as well as the typical work day can help improve accuracy of measurements and discourage underestimates or overestimates of judicial need.

The method we propose offers a “snapshot” of a moving image, as the judicial resources a jurisdiction requires may change over time. For example, sudden sharp increases in cases filed from 2010 to 2011, without accompanying increases in judicial resources, are likely to result in significantly larger estimates of judicial resource needs. Although this method is best utilized with a three-year workload average, which may mitigate some of the challenge to measuring judicial workload, some jurisdictions may be more (or less) capable of conducting substantive and thorough hearings than reflected with the method.

CONCLUSION

The workload assessment method presented here can have a critical impact on courts that need to understand better the complex workloads of their juvenile dependency court judges by providing a more accurate estimate of judicial resource needs in terms of best practices. Best practice recommendations cannot be implemented without an eye to how workload and practice are interrelated. Any jurisdiction that strives to improve practice to achieve better outcomes for children and families must consider the judicial workload resources they have available. As juvenile dependency courts pursue the principles of continuous quality improvement, the workload assessment method outlined in this paper should be a cornerstone of any evaluation of their practices, procedures, and policies.
This workload method can help position courts to communicate with state legislatures about their needs to ensure that sufficient resources are allocated to courts. If legislatures want courts to be able to engage in more than minimally sufficient practice, legislatures will have to provide more than minimally sufficient resources. By ensuring adequate judicial resources, jurisdictions can work to improve engagement of parties, increase the quality of hearing discussion, and ultimately work toward achievement of their goal to increase timely permanency for children and families.

REFERENCES


