An Overview of the Courts Catalyzing Change Preliminary Protective Hearing Benchcard Study

By Jesse Russell and Alicia Summers

ABSTRACT

This paper presents an overview of the Courts Catalyzing Change: Achieving Equity and Fairness in Foster Care Preliminary Protective Hearing Benchcard Study. In the fall of 2009, the National Council of Juvenile and Family Court Judges (NCJFCJ) began a study to examine the effects associated with judges’ use of the Preliminary Protective Hearing Benchcard. For this study, data were gathered from case file information (both court and agency files) and from courtroom observations of more than 500 children in Los Angeles, California; Omaha, Nebraska; and Portland, Oregon. Data from a baseline sample were collected at each of the three sites, and judicial officers at each site were randomly assigned to either a Benchcard group or a control group. Benchcard implementation appears to be associated with more discussion and higher quality discussion of key dependency topics during preliminary protective hearings. Benchcard implementation also corresponds to increased judicial inquiry and parental engagement. Benchcard use also was associated with more family placements—placement with a charged parent, a non-charged parent, or a relative—at the initial hearing and even more family placement at adjudication when comparing the same judges before and after Benchcard implementation. Similarly, the percentage of children who were reunified with the charged parent at the initial hearing and the adjudication hearing increased after Benchcard implementation.

This paper presents an overview of the Courts Catalyzing Change: Achieving Equity and Fairness in Foster Care Preliminary Protective Hearing (PPH) Benchcard Study. The National Council of Juvenile and Family Court Judges (NCJFCJ) conducted this study to assess judicial decision making within the Courts Catalyzing Change (CCC) National Agenda for Reducing Racial Disproportionality and Disparities in the Dependency Court System. Included in this paper are a description of the study, research
informing the development of the PPH Benchcard, and highlighted findings concerning improved court practices and related outcomes.

The CCC agenda, supported by Casey Family Programs and the Office of Juvenile Justice and Delinquency Prevention, was created and launched through the NCJFCJ Model Courts project. The goal of the agenda is to reduce disproportionality and disparate treatment, ultimately improving outcomes for all children in care. The CCC agenda includes efforts to: (1) engage national, state, local, and tribal stakeholders, community partners, and children and families; (2) transform judicial practice from the bench; (3) participate in policy and law advocacy; (4) examine and employ research, data, and promising practices; and (5) impact service array and delivery.¹

The PPH Benchcard is one tool developed as part of this agenda, focusing on a reexamination of the Resource Guidelines: Improving Court Practice in Child Abuse and Neglect Cases (NCJFCJ, 1995) through a racial equity lens and on the development of a specific set of judicial decision-making tools. Intended for use during the initial dependency hearing, the Benchcard asks judges to reflect on the decision-making process to identify and attempt to minimize institutional bias and to consider some key inquiries, analyses, and decisions relating to removal, placement, and services. The Benchcard is built around two types of inquiry: internal and external. The internal inquiry is set forth in a self-reflection section containing questions designed to help judges examine potential biases that may affect their decisions. The external inquiry is laid out in the due process related questions and considerations as well as the actual judicial inquiry of the hearing participants related to specific salient issues that should be determined at the preliminary protective hearing.

**CHILD WELFARE BACKGROUND**

More than 2 million families are investigated for child abuse and neglect each year in the United States (U.S. Department of Health & Human Services, 2012a). For the fiscal year 2011, in the United States, about 681,000 children were confirmed unduplicated victims of child maltreatment (U.S. Department of Health & Human Services, 2012a). Over the past decade, there have consistently been more than 500,000 children in foster care in the United States at any given time (Child Trends, 2011). Only recently has the number of children in foster care dropped to 400,540 at the end of the 2011 federal fiscal year, with 47% of these children living in non-relative foster care homes (U.S. Department of Health and Human Services, 2012b).

Recent research has highlighted a number of troubling issues facing children in foster care. These issues include, as compared to the general population, greater incidence of drug and alcohol use (Thompson & Auslander, 2007); significantly higher levels of unemployment (Macomber et al., 2008); higher likelihoods of homelessness (Yen, Hammond, & Kushel, 2009; Zlotnick, 2009); higher incidences of teen pregnancy

(Dworsky & DeCoursey, 2009); poorer educational outcomes (Trout, Hagaman, Casey, Reid, & Epstein, 2008); and more experiences with depression (Blome, Shields, & Verdeick, 2009).

Further studies have investigated the experiences of foster care alumni later in life. These findings suggest that youth who were placed in foster care as children are significantly more likely to commit crimes, drop out of school, receive welfare benefits, have substance abuse problems, or be homeless than children who were not placed in foster care (Clausen, Landsverk, Ganger, Chadwick, & Litrownik, 1998; Courtney, Dworsky, Lee, & Raap, 2010; Courtney, Piliavin, Grogan-Kaylor & Nesmith, 2001; Pecora et al., 2010; Vinnerljung, Sundell, Lofholm, & Humlesjo, 2006). In addition, nearly 20% of young prison inmates and 28% of people who are homeless have spent some time in foster care as a youth (Burt et al., 1999).

Additionally, evidence shows that children in out-of-home care are clearly at greater risk for short- and long-term school failure than are other children (Pecora et al., 2006). Research has shown that disruptions relating to foster care can lead to a lack of continuity in schooling (Courtney, Roderick, Smithgall, Gladden, & Nagaok, 2004; Malmgren & Meisel, 2002) such as truancy, grade retentions, and multiple placements (Trout et al., 2008). Further, disruptions relating to foster care are related to slow academic achievement (Courtney et al., 2004), negative educational outcomes, and elevated rates of disability (Trout et al., 2008).

Much of the difficulties these youth and young adults face may be due to maltreatment they have experienced and not to their experience in foster care per se (Samuels, 2011), but research has demonstrated that foster care itself can have negative effects for some children. Research has demonstrated that although abusive and neglectful family environments are undoubtedly harmful to children, removing a child from home may also be traumatic (Lawrence, Carlson, & Egeland, 2006).

The evidence suggests that children on the margin of placement—when stakeholders disagree whether the child should be placed in out-of-home care—tend to have better outcomes when they remain at home, especially older children (Doyle, 2007). Moreover, research has indicated that among children on the margin of placement, children who spend time in foster care have arrest, conviction, and imprisonment rates as adults that are three times higher than those of children who remained at home (Doyle, 2008). Further, an instrumental-variables approach2 to this analysis has suggested that foster care placement leads to an increased likelihood of later delinquency and emergency healthcare episodes (Doyle, in press).

Within this context, research has demonstrated that racial disparities exist within the juvenile dependency system concerning the reported allegations, specifically against African Americans (Hill, 2004; U.S. Government Accountability Office, 2007). Before a case reaches court, African-American children are more likely than other children to be

---

2 The instrumental variables approach mimics a natural experiment. That is, when interested in the effect of being in a certain condition (e.g., foster care), but random assignment to conditions is impossible or impractical, this statistical approach estimates how the assignment of a case to different groups affects the likelihood of being in each condition and then how those likelihoods relate to the outcomes effect (e.g., entering or not entering into the juvenile justice system).
referred to protective services (Gryzlak, Wells, & Johnson, 2005; Magruder & Shaw, 2008; Sabol, Coulton, & Polousky, 2004), and to have allegations of abuse and neglect substantiated (Ards, Myers, Malkis, Sugrue, & Zhou, 2003; Fluke, Yuan, Hederson, & Curtis, 2003). Further, African-American children are more likely to be removed from their homes than children of other racial or ethnic backgrounds (Lu et al., 2004; Magruder & Shaw, 2008).

Theoretical and empirical discussions of disproportionality and disparity are often fraught with misunderstanding (Russell, 2011), but empirical evidence suggests that racial differences may stem from a confluence of (a) a greater need for services, (b) structural racism or implicit biases, and (c) community conditions and supports. Children of color may have greater service needs due to higher poverty and other risk factors (Drake & Jonson-Reid, 2011; Putnam-Hornstein & Needell, 2011; Wulczyn, Barth, Yuan, Jones-Harden, & Landsverk, 1995). Families of color as a whole are more likely to come to the attention of the child welfare system because they are impacted more profoundly by outside factors that can increase the risk of child neglect, such as poverty, joblessness, substance abuse, and mental illness.

There also may be racial bias in the child welfare system (Hill, 2004; Johnson, 2007; McCrory, Ayers-Lopez, & Green, 2006). Some child welfare systems have biased or culturally insensitive practices and policies that can lead to inequitable decision-making processes (Bent-Goodley, 2003; McRoy, 2004; Morton, 1999; Roberts, 2007). These biases and insensitivities may be due to systemic factors or to individual worker biases.

Community conditions and supports may also be inadequate. Geographic contexts and variability may therefore have a disparate impact on communities and families of color in terms of service availability and access (Roberts, 2007; Wulczyn, 2011). Child welfare systems and processes may disparately involve families of color (Courtney, et al., 1996; Fluke, Jones-Harden, Jenkins, & Ruehrdanz, 2010; Texas Health and Human Services Commission and Department of Family and Protective Services, 2006).

Given this myriad of potential problems, the purpose of this study is to assess the Benchcard as a potential tool for ameliorating disparities and negative outcomes in juvenile dependency cases. This intervention does not purport to address every potential source of decision making or disparity; rather it focuses on the role of judicial decision making in juvenile dependency case hearings.

OVERVIEW OF STUDY

In the fall of 2009, NCJFCJ’s Permanency Planning for Children Department began a study to examine the effects associated with judges’ use of the PPH Benchcard. The first phase of the study included data collection on more than 500 children in Los Angeles, California; Omaha, Nebraska; and Portland, Oregon. Data were gathered from case file information (both court and agency files) and from courtroom observations. A longitudinal study was designed to follow these children throughout the course of their involvement in the juvenile dependency system. Researchers from the NCJFCJ, the University of Nevada, and the University of Nebraska-Lincoln collected data at several
junctures, from placement to establishment of jurisdiction and disposition, completing data entry in March 2010. To explore Benchcard implementation effects, the study was designed to allow for several different comparisons. Researchers collected information on numerous data points, including demographic details (including race), information about the families involved, hearing participants, dates of case events, as well as details on allegations, services, and placement. Data from a baseline sample were collected at each of the three sites, and judicial officers at each site were randomly assigned to either a Benchcard group or a control group.

For data collected by courtroom observation, a standardized hearing observation instrument was constructed. The instrument included items related to the content and level of discussion in the hearing, the length of the hearing, parties present, and parent engagement. Researchers were trained on the instrument during a two-day session (which also included training on the case file review instrument). During the training, all researchers discussed the instrument and considered any potential for coding discrepancies. Researchers visited each of the three study sites during the first week of Benchcard implementation. In Portland and Omaha, researchers attended every preliminary protective hearing that was held during a four-day site visit. Due to differences in caseload volume in Los Angeles, each researcher was assigned to a block of courtrooms for one day. Researchers then coded the first three hearings they observed in pairs. Coding was then compared to check-code and to generate inter-rater reliability scores (which indicated no significant or substantive differences). Afterwards, researchers attended hearings individually for the remainder of the day.

Judicial officers in the Benchcard group were trained on its use, including receipt of a draft NCJFCJ Technical Assistance Bulletin explaining the development of the Benchcard (Right from the Start: A Judicial Tool for Critical Analysis & Decision-Making at the PPH Hearing). They began implementing the Benchcard in their preliminary protective hearings. Each randomly assigned judicial officer heard 10 preliminary protective hearings using the Benchcard, while the control group of judicial officers in each of the sites heard 10 preliminary protective hearings without Benchcard implementation. To isolate the judicial intervention, the Benchcard was not shared with stakeholders during the research project. Had the Benchcard been shared with others, it would have been difficult to measure whether judges or others were raising key issues associated with Benchcard use.

Four research questions are posed:

- Research Question 1: How does Benchcard implementation affect the thoroughness of preliminary protective hearings?
- Research Question 2: Does Benchcard implementation affect the rates that children are placed with family (including with parents) at the preliminary protective hearing and at the adjudication hearing?
- Research Question 3: Does Benchcard implementation affect the likelihood that a child will be returned home at the preliminary protective hearing?
- Research Question 4: Do the numbers and types of allegations vary substantially across racial groups?
METHOD

Design

The study design included a quasi-experimental and an experimental design. The quasi-experimental design consisted of a pre/post test comparing the decisions of Benchcard users before and after implementing the Benchcard. In the experimental design, judges from each site were randomly assigned to implement the Benchcard or not (as a control group). Judges were trained on the use of the Benchcard and asked to implement the Benchcard into daily practice at the preliminary protective hearing.

Before implementing the Benchcard, each study site was required to develop or enhance an existing multi-system collaborative group to guide local implementation of the Courts Catalyzing Change agenda. The courts were then required to participate in training about institutional and structural racism and implicit bias to better understand the history and complexity of these important issues. The courts also developed strategic plans for CCC implementation. (See Model Courts National Agenda Implementation Guide at http://www.ncjfcj.org/sites/default/files/MC%20National%20Agenda.pdf).

Participants and Sites

This study included a review of 555 cases, with the number of observations varying by point of comparison. Of the 555 cases, 336 were from the pre-implementation baseline, and 219 were from post-implementation. Also, of the 555 cases, 277 were from Benchcard users (both pre- and post-implementation), and 278 were from the control group (both pre- and post-implementation). Of the 219 post-implementation cases, 119 were Benchcard users and 100 were in the control group. Sample size by treatment condition is presented in Table 1 and summary statistics of sample characteristics are presented in Table 2.

Because of the complex nature of child abuse and neglect hearings, differences in practice, policy, and statutes may occur across jurisdictions. The three project sites were selected in part because they do exhibit some diversity in their approaches. They are

### TABLE 1
Study Participants

<table>
<thead>
<tr>
<th>Phase</th>
<th>Treatment Group</th>
<th>Control Group</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-Implementation</td>
<td>158</td>
<td>178</td>
<td>336</td>
</tr>
<tr>
<td>Post-Implementation</td>
<td>119</td>
<td>100</td>
<td>219</td>
</tr>
<tr>
<td>Total</td>
<td>277</td>
<td>278</td>
<td>555</td>
</tr>
</tbody>
</table>

3 True randomization at every site did not occur as anticipated. In one study site, a judge randomly assigned to the control condition had been involved with the Benchcard construction and thus had to be moved to the Benchcard condition. In another site, the sample for the Benchcard group was drawn only from judges who were willing and able to participate in the study, not the entire sample of judges.
Similar in many regards, but there were some notable differences among the three jurisdictions.

Specifically, in Omaha, new cases come to a pre-hearing conference (Douglas County District Court Rule 4-15). The focus is on alternative dispute resolution, during which all parties (except the judge) meet and discuss issues relevant to the preliminary protective hearing. Following this pre-hearing conference, the judge holds a preliminary

<table>
<thead>
<tr>
<th>Primary Race or Ethnicity of Child</th>
<th>Control Group</th>
<th>Benchcard Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>White/Caucasian</td>
<td>35%</td>
<td>35%</td>
</tr>
<tr>
<td>Black/African American</td>
<td>22%</td>
<td>24%</td>
</tr>
<tr>
<td>Latino/Hispanic</td>
<td>29%</td>
<td>29%</td>
</tr>
<tr>
<td>Native American/Indian</td>
<td>0%</td>
<td>2%</td>
</tr>
<tr>
<td>Asian/Pacific Islander</td>
<td>3%</td>
<td>1%</td>
</tr>
<tr>
<td>Other race or ethnicity</td>
<td>4%</td>
<td>4%</td>
</tr>
<tr>
<td>Unable to determine</td>
<td>7%</td>
<td>7%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Jurisdiction</th>
<th>Control Group</th>
<th>Benchcard Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Omaha</td>
<td>24%</td>
<td>30%</td>
</tr>
<tr>
<td>Los Angeles</td>
<td>45%</td>
<td>39%</td>
</tr>
<tr>
<td>Portland</td>
<td>30%</td>
<td>31%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gender</th>
<th>Control Group</th>
<th>Benchcard Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>41%</td>
<td>39%</td>
</tr>
<tr>
<td>Female</td>
<td>54%</td>
<td>54%</td>
</tr>
<tr>
<td>Unable to determine</td>
<td>5%</td>
<td>7%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Age</th>
<th>Control Group</th>
<th>Benchcard Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>3 or under</td>
<td>36%</td>
<td>38%</td>
</tr>
<tr>
<td>3.1 thru 9</td>
<td>28%</td>
<td>23%</td>
</tr>
<tr>
<td>9.1 thru 13</td>
<td>11%</td>
<td>9%</td>
</tr>
<tr>
<td>13.1 thru 18</td>
<td>18%</td>
<td>18%</td>
</tr>
<tr>
<td>Unable to determine</td>
<td>8%</td>
<td>12%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Parent Status</th>
<th>Control Group</th>
<th>Benchcard Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mother is single</td>
<td>50%</td>
<td>54%</td>
</tr>
<tr>
<td>Mother is not single</td>
<td>32%</td>
<td>29%</td>
</tr>
<tr>
<td>Father is single</td>
<td>1%</td>
<td>0%</td>
</tr>
<tr>
<td>Unable to determine</td>
<td>17%</td>
<td>17%</td>
</tr>
</tbody>
</table>

T-tests were run for differences in each of the above factors between the control group and the Benchcard group. The tests showed these differences to be statistically insignificant at the 0.05 level for each of the above factors with the exception of Asian/Pacific Islander. However, the scale of the difference between the control group and Benchcard group on this factor is too small to be substantive.

**TABLE 2**
Characteristics of the Sample
(Total number of observations = 555)
protective hearing. In Los Angeles, judges routinely have higher caseloads than judges in
the other two sites. California also routinely appoints counsel for all parties at the first
hearing (CA Welf. & Inst. Code §317; Cal. Rules of Ct., Rule 5.660) and may expedite
the processing of special cases (e.g., children under the age of 3 years). In Portland, an
initial hearing is held immediately following the petition filing, and a second expanded
preliminary hearing may be set within 30 days to gather more information (Supplemen-
tal Local Rules, Fourth Judicial District, Circuit Court of the State of Oregon for Mul
tnomah County 11.045-2a); counsel is appointed for parents at the first hearing
(ORS § 419B.205) if they are present in court.

Despite these differences, these sites are similar in the way that they handle child
abuse and neglect cases. The sites not only follow the federal mandates for child abuse and
neglect case processing (e.g., the Adoption and Safe Families Act), but they also are all
NCJFCJ Model Court4 sites, dedicated to systems change and implementation of best or
promising practices.

Measures

There were four dependent variables of direct interest: (1) thoroughness of prelimi-
nary protective hearings; (2) relative placement decisions made at the preliminary pro-
tective hearing, and placement decisions made at the adjudication hearing; (3) whether
a child is placed at home at the conclusion of the preliminary protective hearing; and
(4) allegations.

Hearing Thoroughness

Researchers rated hearing thoroughness according to a standardized observation
metric on the amount and breadth of discussion on each topic (ranging from 1 = not
discussed at all, to 4 = substantively discussed). Topics included on the observation
metric included: discussion of the petition, parties who should be present, services offered
and their appropriateness, placement of the child, whether the child can go home today,
and reasonable efforts (active efforts in Indian Child Welfare Act cases; 25 U.S.C. §
1912). The topics also included discussion of the allegations and identification of parties,
any paternity issues, probable cause for removal, cultural and linguistic issues relating to
removal and placement, offering services that might allow the child to go home, safety
planning, engaging parents in developing services, and focusing on strengths.

Judicial inquiry was coded on a scale from 0 to 1 for each topic. Parental engage-
ment for each group was coded on a scale from 0 = no engagement to 1 = substantial
engagement (based on a direct judicial inquiry, parental opportunities to speak, to ask
questions, or to indicate understanding). Weighted scores were created by multiplying
the number of items discussed with the thoroughness of the discussion.

4 Model Courts typically consist of a lead judge (or judges), who works closely with an interdisci-
plinary collaborative team of system stakeholders, in efforts to implement best practices and evoke systems
change. More information on the NCJFCJ Model Courts can be found at http://www.ncjfcj.org/our-work/
dependency-model-courts.
Whether a child was returned home was coded for children who were placed in out-of-home care up to the preliminary protective hearing, and were then placed with a parent at the conclusion of the preliminary protective hearing.

Placement Decisions

Placement decisions were coded by researchers based on the court orders for the case. Family placements included placement with a charged parent, with a non-charged parent, or with a relative.

Allegations

Allegations were coded based on those listed on the petition. Allegations include items such as physical or sexual abuse, neglect, threat of harm, and failure to protect. Allegations may also include presenting problems, such as substance abuse (drugs or alcohol) or poor mental health functioning.

Researchers collected case file review data on a small sample of cases the first week immediately following implementation of the Benchcard into daily practice, and also collected case file review data on the first 10 new cases immediately following the scheduled implementation of the Benchcard. The selection of 10 cases was meant to help reduce the potential for spillover effects, as some attorneys and caseworkers work with multiple judges and may begin to respond to Benchcard questions even if in front of a control group judge who does not have the Benchcard. Judges were not given an “end date” to stop using the Benchcard. Researchers also observed and coded a sample of court hearings immediately following implementation of the Benchcard. For the court observation, researchers used a structured instrument that was designed to assess fidelity to the Benchcard implementation. That is, the instrument measured thoroughness of discussion, as well as judicial inquiry into key topic areas.

RESULTS

Hearing Thoroughness

Based upon systematic courtroom observations, a rigorous count of the items discussed indicates that judicial officers who used the Benchcard discussed more key topics during the preliminary protective hearings than did the control group. Benchcard implementation appears to be associated with substantially more discussion and higher quality discussion of key dependency topics when compared to the control group. In addition, t-tests indicate that these differences are statistically significant at the \( p = 0.05 \) level (\( n = 29 \)).
Benchcard implementation also corresponds to an increased thoroughness of discussion and judicial inquiry. In addition to how many topics parties discussed, there were differences in how thorough that discussion was. Thoroughness of discussion was assessed in terms of both the number of topics covered and the thoroughness of discussion into each of those topics, in terms of topics outlined in the Resource Guidelines and those framed by the CCC initiative.

Similarly, how much the judicial officer inquired into the topics and how directly the judicial officer engaged the parents in the discussion differed between the Benchcard group and the control group. Judicial officers in the Benchcard group engaged parents more than judicial officers in the control group. Again, t-tests indicated that these differences were statistically significant at the 0.05 level.

These results were similarly true when assessing judges in the Benchcard group before and after implementation. Across the three sites, the percentage of Resource Guidelines topics discussed by Benchcard users pre-implementation was 37% compared to 76% post-implementation. Also, the percentage of CCC topics discussed by Benchcard users pre-implementation was 32%, compared to 56% post-implementation. A t-test indicated that these pre/post differences are statistically significant at the 0.05 level.

These process findings vary substantially across the three study sites.

- In terms of the percentage of Resource Guidelines topics discussed at the initial hearing, Portland was the lowest of the three sites in the control group, but it reached 100% of topics discussed in the Benchcard implementation group. Omaha was highest in the control group, and the Benchcard implementation group was not much higher.
- For the percentage of CCC topics discussed, Portland was at the highest level of the three sites in the control group, and it was also at the highest level of the three sites in the Benchcard implementation group. Los Angeles’ control group was at the lowest level of the three, and though it was increased in the Benchcard implementation group, it was at the lowest of the three in terms of CCC topics discussed during the initial hearing.
- For parental engagement, Portland’s control group was at the lowest level of the three, and it was also at the lowest level of the three in the Benchcard implementation group, but the difference there was markedly greater than in the other two sites. Parental engagement in Omaha was at a high level in the control group, and it actually was found to be lower in the Benchcard implementation group (though the difference was nominal).

Overall, these findings indicate that experiences with Benchcard implementation varied. Baseline levels of Resource Guidelines and CCC practice were widely different, and Benchcard implementation differed notably across sites.

Assessing case outcomes, a comparison of the same judges before and after Benchcard implementation found that a higher percentage of children were returned home after
the initial hearing with Benchcard implementation. Before Benchcard implementation, Los Angeles hearings were associated with more children returning home after the initial hearing than were hearings in Portland and Omaha. With Benchcard implementation, Portland and Los Angeles increased the number of children returned home after the initial hearing, with Portland increasing more. For Portland, the effect was statistically significant at the 0.05 level (based on a logistic regression), but the difference was not statistically significant in Los Angeles. In Omaha, the number of children returning home decreased marginally with Benchcard implementation.

*Placement Decisions*

Benchcard use was related to more family placements—placement with a charged parent, a non-charged parent, or a relative—at the initial hearing and even more placements at adjudication when comparing the same judges before and after Benchcard implementation. Reciprocally, Benchcard use was also associated with fewer children being placed in non-relative foster care at the initial hearing and even fewer at adjudication. A logistic regression comparing the likelihood of a family placement versus non-family placement before and after Benchcard implementation shows these findings to be statistically significant at the 0.05 level. Substantial differences did exist across the three sites, but the findings remained significant when the three sites were accounted for in the statistical model.

Similarly, the percentage of children who were reunified with the charged parent at the initial hearing and the adjudication hearing increased after Benchcard implementation. A logistic regression shows the likelihood of reunification to be different at a statistically significant (0.05) level between before and after Benchcard implementation. Again, substantial differences did exist across the three sites, but the results were significant when sites were accounted for in the statistical model.

*Allegations*

The study found race differences in filing trends. White mothers (in comparison to African-American and Hispanic mothers) entered court with a higher number of allegations, averaging 3.3 allegations compared to 2.4 for others (ANOVA analysis shows this difference to be statistically significant at the 0.05 level). White mothers also had more allegations of substance abuse, homelessness, and mental health issues (each of these represents statistically significant differences at the 0.05 level). Overall, African-American mothers came into court with fewer allegations and were more likely to have their case dismissed. African-American mothers tended to be brought to court with less serious allegations. Hispanic mothers were brought to court with the fewest number of allegations on average, and they had the least serious allegations.

Differences were found by racial group in “failure to protect” allegations, which can mean that no maltreatment has yet occurred but the risk of maltreatment was present. This type of allegation is often seen in conjunction with other allegations of neglect or abuse or in domestic violence cases in which the mother is the victim of the violence.
White families were substantially less likely to face this allegation than families of other racial groups. Conversely, Hispanic families were substantially more likely to face an allegation of failure to protect.

Allegations of substance abuse show a different pattern. White families in the sample were much more likely to face allegations of failure to supervise or parent adequately due to substance abuse (drugs or alcohol) than were families of other racial groups (again an ANOVA shows these differences to be statistically significant at the 0.05 level). Allegations of substance abuse for African-American and Hispanic families were somewhat less common.

Differences among racial groups are also apparent in allegations involving poor parenting due, in major part, to poor mental health functioning. As with substance abuse allegations, White families were much more likely to be brought to court with allegations relating to mental health. White families were almost three times as likely to face a mental health allegation as families from other racial groups (again statistically significant at the 0.05 level).

**DISCUSSION**

The empirical results indicate that judicial officers who used the Benchcard discussed more key topics during the preliminary protective hearings than did the control group. Benchcard implementation appears to be associated with more discussion and higher quality discussion of key dependency topics identified in both the *Resource Guidelines* and the CCC initiative when compared to the control group. Benchcard implementation also corresponds to an increased thoroughness of discussion and judicial inquiry, as demonstrated by the number of topics and how thoroughly they were discussed. These process findings indicate that Benchcard implementation is associated with substantial increases in the quantity and quality of discussion in juvenile dependency hearings.

Benchcard use also was associated with more family placements—placement with a charged parent, a non-charged parent, or a relative—at the initial hearing and even more again at adjudication when comparing the same judges before and after Benchcard implementation. (Reciprocally, Benchcard use was also associated with fewer children placed in non-relative foster care at the initial hearing and even fewer again at adjudication.) Similarly, the percentage of children who were reunified with the charged parent at the initial hearing and at the adjudication hearing increased after Benchcard implementation. Differences did exist across the three sites, but the findings remained significant when the three sites were accounted for in the statistical analysis.

As with all experimental research, the present study is limited. All three study sites were Model Courts as part of NCJFCJ’s Model Court Project, and thus it may be inappropriate to generalize beyond them. Further testing of the Benchcard is expected to include other court jurisdictions. Also, given the design of this study, findings reflect only what occurred once a case was brought to court, not what occurred prior. For example, the study did not collect information on decisions around which cases were reported to child protective services, were investigated, or were substantiated.
The fact there may have been observational effects should also be considered. The judges were informed when the research began and may have modified their behavior because they were being watched. However, any observational effects are likely to have been mitigated for three reasons. First, all three sites are Model Court sites, which NCJFCJ staff regularly observe. Second, though some of the participating judges were observed in court, many were not. Also, participating judges were not reminded of the study and, as is often the case with observation, may have quickly returned to “typical” practice.

Some of the empirical findings presented in this paper do not have obvious causal explanations. The finding that White mothers face more allegations and more serious allegations than African-American mothers, for example, may be because White mothers have a higher threshold to cross in terms of case allegations before being brought to court. This hypothesis, however, is not testable with this study’s data. Similarly, the finding that more services are ordered for White mothers than for other mothers may simply represent a greater tendency to offer services to White mothers, but may also indicate differences in the seriousness or number of allegations by the time White mothers’ cases are brought to court. Again, further research would be needed to explore these ideas.

The major limitation of this analysis was the potential for “spillover” effects. Once judges began asking questions of social workers and attorneys, these professionals began preparing these answers for court, regardless of the judge. This preparation means that many of the Benchcard question responses were being given to control group judges. Some jurisdictions use teams per courtroom or per judge, which may reduce this effect. To reduce the potential for spillover effects further, the research team used only the first 10 new cases from each judge, thus reducing the amount of spillover due to the decreased time for the spillover to occur. In addition, there was a decrease in sample size for the post-implementation cases, thereby reducing the likelihood of finding statistically significant differences.

CONCLUSION

Judicial decisions are made every day regarding temporary or permanent removal of children from their families. Judges must be aware of factors that may influence their decision making. In this study, it is important to note that the foundational training provided before Benchcard implementation is an important component of its implementation. Tools have been developed over many years for judges and other stakeholders by many esteemed organizations and individuals. These tools are implemented absent an examination of how an individual’s own bias may affect his or her effectiveness. In the Courts Catalyzing Change initiative, judges made a commitment to explore their own personal decision making and examine the possible effects of implicit bias and structural and institutional racism.

Congress and state legislatures have vested the courts with oversight responsibility of the child welfare system. Both federal and state laws require that foster care be utilized only as a last resort. When judges conduct thorough hearings, they are more able to
determine whether removal decisions are truly last-resort decisions. They are also more likely to be able to identify relative resources so children can maintain their connections with family.

This study highlights the potential for the Courts Catalyzing Change Preliminary Protective Hearing Benchcard to improve placement outcomes (less foster care, more relative care) in juvenile dependency cases. With more than 400,000 children in foster care, the placement improvements associated with Benchcard implementation deserve significant notice. Research shows that removing children from their families and homes can create trauma and that foster care experiences can be deleterious to the long-term well-being of some children. As courts aim to place more children with family and safely reunite more children with their families sooner, the Benchcard can be an important tool for system improvement. Implemented by a strong judicially led, multidisciplinary collaborative, the Courts Catalyzing Change initiative has the potential to affect the lives of thousands of children and families by returning children home and maintaining family connections.

REFERENCES


