
OVERVIEW: FEDERAL PERFORMANCE MEASUREMENT SYSTEM

The Context of Child Welfare Performance Measures

Sarah Carnochan, Mark Samples, Jennifer Lawson, and Michael J. Austin
*Bay Area Social Services Consortium, School of Social Welfare, University of California,
Berkeley, California, USA*

A growing movement in the United States seeks to improve the outcomes achieved by public agencies through performance measurement strategies. In child welfare policy, a federally mandated performance measurement system has evolved since the 1980s, establishing the federal Child and Family Services Review which uses a series of performance indicators for evaluating the child welfare system. This article reviews the literature on performance measurement and performance management in the public sector in order to develop a conceptual framework for examining the federal child welfare performance measurement system. It briefly summarizes the evolution of federal policy related to performance measurement in child welfare. The framework is then used to guide an analysis of the debate surrounding the establishment of the current child welfare performance measurement system, concluding with recommendations for consideration in future reform efforts.

Keywords: Child welfare, performance measurement, policy, accountability

Since the 1960s, there has been a growing movement in the United States to improve the outcomes achieved by public agencies through strategies of performance management (Heinrich, 2002). The central focus has been developing performance measurement systems aimed at increasing accountability, improving service quality, and achieving better outcomes for individuals served (Heinrich, 2002; Moynihan & Pandey, 2005; Schalock & Bonham, 2003). As Heinrich argues, the increasing level of autonomy and discretion granted to state and county agencies, under policies such as the block grant programs of the Reagan era, raised concerns about ensuring accountability, providing the “impetus for the introduction of new mechanisms for performance accountability” (2002, p. 713).

Address correspondence to Sarah Carnochan, University of California, Berkeley, Mack Center for Nonprofit Management in the Human Services, School of Social Welfare, 120 Haviland Hall #7400, Berkeley, CA 94720-7400, USA. E-mail: scarnochan@berkeley.edu

Beyond the direct potential to improve services and outcomes for those served, additional uses and benefits have been identified, including to: “target effective services for expansion; develop and justify budgets; prepare long-range plans, engage collaborators; retain and increase funding; and gain favorable public recognition” (Wells & Johnson, 2001, p. 193, citing Hatry, van Houten, Plantz & Greenway, 1996). Scholars have noted the benefits associated with performance-based systems, including the potential “to support efficient and effective services utilization, more accurately identify service needs and communicate them to the public, and measurably increase the effectiveness and efficiency of the child welfare system” (Wells & Johnson, 2001, p. 171).

In child welfare policy, a federally mandated performance measurement system has evolved since the 1980s, establishing the federal Child and Family Services Review (CFSR) which incorporates a series of specific performance measures for evaluating the child welfare system. Child welfare professionals have demonstrated substantial agreement on the overall intent of the CFSR process, but compliance with the standards defined by the performance measures has posed clear challenges. While a number of state officials have pointed to benefits resulting from participating in the CFSR process (e.g., improved collaboration with community stakeholders and partners), some have questioned the accuracy of the administrative data and the efficacy of imposing financial penalties to achieve state and county conformity with the performance measures (General Accounting Office [GAO], 2004). Other experts have raised a number of concerns about the specific performance measures being utilized (Schuerman & Needell, 2009).

In this article the authors review the literature on performance measurement and performance management in the public sector, in order to develop a conceptual framework for examining the federal child welfare performance measurement system. They then briefly summarize the evolution of federal policy related to performance measurement in child welfare. The framework is then used to guide an analysis of the debate surrounding the establishment of the current child welfare performance measurement system, concluding with recommendations for consideration in future reform efforts.

A CONCEPTUAL FRAMEWORK FOR PERFORMANCE WITHIN CHILD WELFARE

The inter-relationship between the development of performance measures and the use of performance management within public human service organizations is receiving more attention in the literature (Abma & Noordegraaf, 2003; Bovens, Schillemans & Hart, 2008; Kaplan, 2001; Perrin, 1998; Speckbacher, 2003; Van Thiel & Leeuw, 2002). The development of public sector performance measures often originates with benchmarked standards created by legislative and regulatory bodies external to human service organizations (Heinrich, 2007; Moxham & Boaden, 2007; Woods & Grubnic, 2008). However, multiple stakeholders play a role in the development of performance measures by sharing different perceptions of success related to an organization’s mission and service goals (Broad, Goddard & Alberti, 2007; Forbes, 1998; Kaplan, 2001; Speckbacher, 2003; Yetano, 2009). This process includes external stakeholders at the national, state, and local level who provide both legitimacy and funding for public agencies along with multiple local stakeholders who help define performance outcome measures (Melkers & Willoghby, 2005; Sanger, 2008; Moynihan & Pandey, 2005).

Developers of performance measures seek to balance the perceptions of all these stakeholders in order to avoid creating multiple, ambiguous, and conflicting measures of effectiveness (Abma & Noordegraaf, 2003; Perrin, 1998; Van Thiel & Leeuw, 2002). However, balancing these perceptions is also influenced by the power of external stakeholders who control fiscal resources to impose their values and beliefs on the process (Abma & Noordegraaf, 2003; Behn, 2002; Benjamin, 2008;

Ebrahim, 2005; Hatry, 2002). Balancing stakeholder perceptions not only affects the development of performance measures but also their utilization in agency performance management.

Performance management practices are generally internal organizational processes incorporating logic models, program evaluation, and strategic planning in order to create “results-oriented systems” (Hatry, 2006; Melkers & Willoghby, 2005; Spechbaker, 2003). In this context, performance management represents a decision-making process that seeks to balance local internal stakeholder experiences (from top administrators to line staff) with externally-defined accountability metrics (Sanger, 2008; see also Woods & Grubnic, 2008). Sanger (2008) notes that effective performance management systems are built upon: (1) nurturing local stakeholder involvement in the process; (2) creating goals that are specific and logically linked to metrics that measure progress toward those goals; and (3) continually fine-tuning measures and goals that are strategically linked to balancing the needs of federal and state funders with those of clients and local citizens.

Linking performance data for external accountability with internal agency evaluation and learning for the purpose of agency-level decision making is difficult in practice. Among the potentially negative consequences of maintaining this balance is the *performance paradox*, in which external accountability designed to improve outcomes results in agency responses that either have no effect on true outcomes, or, in the worst cases, decrease service quality and lead to more negative outcomes (Van Thiel & Leeuw, 2002). For example, public agencies face the temptation to engage in results-oriented decision making that is more focused on manipulating a political process among legitimizing external stakeholders through the use of various compliance activities than objectively assessing outcomes of direct importance to local agency stakeholders (Abma & Noordegraaf, 2003; Dahler-Larsen, 2007; Poertner, Moore & McDonald, 2008; Salaman & Storey, 2002).

In contrast, proponents of performance measurement and accountability in the public sector have worked to strengthen and refine the conceptualization of performance measures and develop specific tools to improve their implementation in practice (Heinrich, 2002; Moynihan & Pandey, 2004; Schalock & Bonham, 2003). The movement to develop and implement internal agency performance assessment and learning is founded on the assumption that “management matters to performance and effectiveness” (Moynihan & Pandey, 2004, p. 422).

The level of influence held by the different stakeholders participating in the development of performance measures and in agency performance management processes is shaped and constrained by a range of contextual factors. Generally, human service agencies operate within their own unique political economy, and are subjected to limited and fluctuating financial resources (Fellin, 2001; Hasenfeld, 1983). Political, cultural, and regulatory forces also affect the development of performance measures and the impact of performance management systems (Dahler-Larsen, 2007). For instance, state legislatures vary in their approaches to the federal legislative mandates for the development and regulation of performance measurement systems in child welfare, which, in turn affects the way that local agencies interpret and act upon performance standards (Schuerman & Needell, 2009).

The environmental context of child welfare services also includes the demographic characteristics of children, their biological families and kin, as well as foster parents and adoptive families. For instance, children who have been in the child welfare system longer, older children, children with siblings, children with special needs, and children of specific racial backgrounds all represent cases in which timelines for permanency are affected (Albers, Reilly & Rittner, 1993; Park & Ryan, 2009; Needell & Putnam-Hornstein, 2009a, 2009b). Similarly, Needell & Putnam-Hornstein (2009a) note a number of factors that may affect state or county performance related to the federal performance measures, including, for example, the relationship between reunification timeliness and the availability of preventative services for families of origin, in addition to child characteristics of race, age, and special needs. Interpreting child welfare performance data through a comparison with national standards or between different states and counties is thus complicated

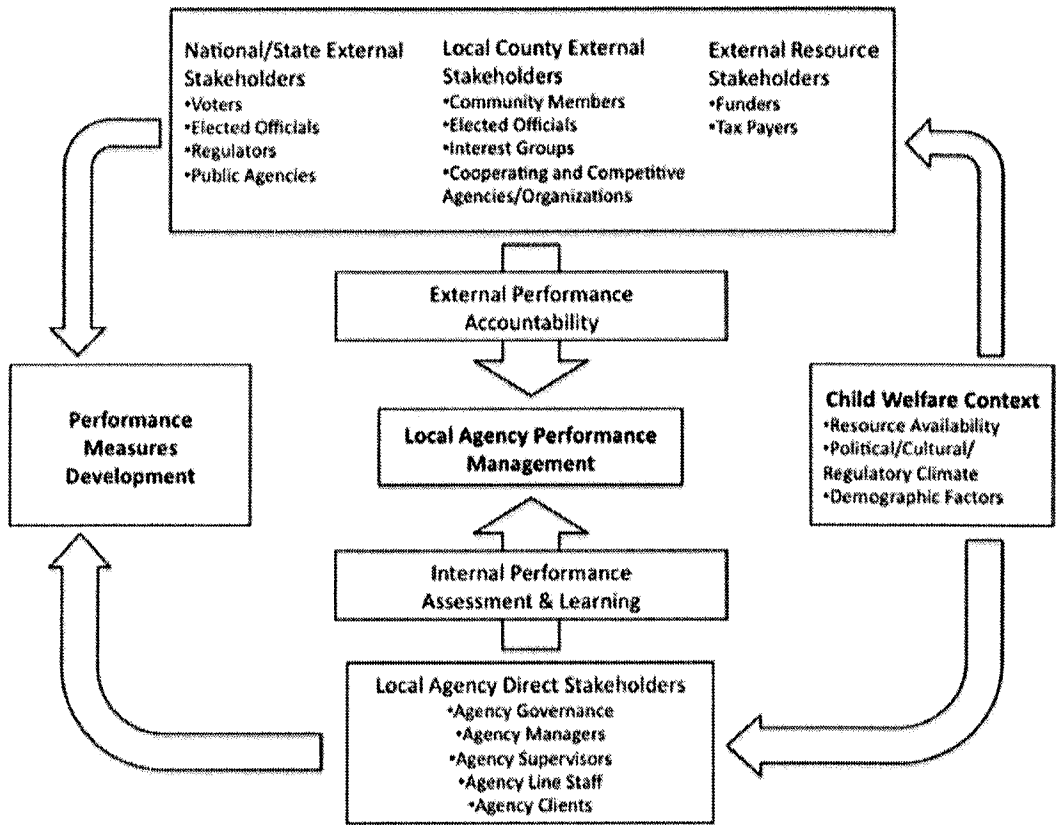


FIGURE 1 A framework for interpreting public agency performance.

by the unique demographic makeup characterizing different regions (Needell & Putnam-Hornstein, 2009a; Schuerman & Needell, 2009).

In this review we will highlight several key concepts relevant to understanding performance measurement systems, namely the importance of including multiple stakeholder perspectives, the role of environmental context in shaping service outcomes, and the complex interaction between external performance measurement goals and local performance management priorities. The relationship between these concepts is illustrated in Figure 1, which will be used to frame the analysis of the federal child welfare performance measurement system. Before turning to the analysis, in the following section we provide a brief summary of the legislative history leading up to the current system.

THE LEGISLATIVE POLICY CONTEXT FOR PERFORMANCE MEASURES

The shifts in federal policy that have governed child welfare policy over time are reflective of the changing attitudes and opinions on the issues surrounding permanency and the role of the government in child protection (Murray & Gesiriech, 2004; Testa, 2008). The current policy focus on permanency in child welfare developed, in part, as a response to the growing awareness that numerous foster children were languishing in the foster system for years, often moving in

and out of multiple placements without returning home or being placed in adoptive homes, a phenomenon known as “foster care drift” (Testa, 2008). This problem led to the development of permanency planning as a policy remedy that was first codified as a federal strategy in 1980 with the passage of the Adoption Assistance and Child Welfare Act (Testa, 2008; Ward, Hamilton, Fein, & Maluccio, 1982). In this political context, permanency emerged as a responsive case planning strategy for reducing the financial costs and social undesirability associated with children lingering in foster care.

Testa (2008) provides a useful analytical framework for looking at child welfare policy, asserting that child welfare services evolve over time based on changing public opinions about the extent to which government should intervene in families and the appropriate strategy for meeting the needs of children when intervention does take place. As Testa (2008) notes:

The key tensions and questions that characterize contemporary child welfare policy and practice may be conceived as arising from two opposing tendencies: the first concerned with the scope of public interest in child welfare and the second concerned with the appropriate form of social organization for meeting these responsibilities. (p. 109)

According to Testa (2008), the public response to child welfare shifts over time according to changing public attitudes on the following two dimensions:

- (1) *Scope of public interest.* Attitudes on this dimension cycle between *narrow* (i.e., the government should only intervene in cases of severe threat to child safety and health) and *diffuse* (i.e., the government should intervene whenever a child’s overall well-being can be advanced).
- (2) *Type of social organization to meet government responsibility.* Attitudes on this dimension range between *primordial* (i.e., the continuity of kinship ties is the most important factor in meeting the needs of children who enter the child welfare system) and *bureaucratic* (i.e., the needs of children in the child welfare system should be met by the most qualified caregiver, regardless of kinship, birth affiliation, or cultural affinity).

This framework provides a context for understanding the historical development of child welfare policy strategies. The following legislative timeline summarizes the major federal laws that are relevant to the formation of the current federal child welfare system and its strategies for protecting children. The shifting public attitudes identified by Testa (2008) can be seen in the policy changes over time where the more bureaucratic view of government responsibility can be seen in the evolution of child welfare policy to emphasize permanency rather than historical concerns about family preservation and reunification. The next section provides a summary of the timeline of evolving child welfare policy by describing the major federal legislation relevant to an array of performance measures.

Child Abuse Prevention and Treatment Act of 1974

The *Child Abuse Prevention and Treatment Act of 1974* (CAPTA) was the first major federal policy to provide mandated national guidelines for state and local child welfare agencies to put child maltreatment reporting and investigation services in place supported by federal funding for states to provide child welfare prevention and intervention services (Murray & Gesiriech, 2004; Pecora, 2006). The passage of CAPTA in 1974 led to a rapid expansion of the number of children in foster care throughout the country due to the mandated reporting systems for child abuse and neglect allegations (Murray & Gesiriech, 2004).

could occur before other permanency arrangements (such as adoption or kinship guardianship) were implemented (Adler, 2001; Berrick, 2009). ASFA emphasized adoption as part of permanency planning by initiating the following changes in child welfare policy and practice (Adler, 2001; Fox, Frasch, & Berrick, 2000; Kemp & Bodonyi, 2002; Reed & Karpilow, 2009):

- (1) Stricter limits on the time children spend in foster care, mandating that a permanency hearing be held for any child still in foster care after 12 months in order to plan for alternative permanency arrangements if reunification is no longer plausible.
- (2) Required the termination of parental rights if a child has been in foster care for 15 of the most recent 22 months.
- (3) Established a set of child welfare outcome measures in order to facilitate a systematic way of collecting data and measuring the progress made by states in achieving results.
- (4) Promoted adoption as a permanency strategy by providing incentive funds to states for increasing adoptions of children out of foster care.

ASFA remains the central federal policy governing the operations of child welfare agencies today. Its emphasis on evaluation and measurement of agency performance set the course for the development of the current federal child welfare evaluation process, the Child and Family Services Review (CFSR).

CHILD AND FAMILY SERVICES REVIEWS

The current CFSR process was initiated in 2001, based on federal mandates for child welfare performance evaluation that were codified in the Social Security Amendments of 1994 and ASFA of 1997 (GAO, 2004; CFR, 2007). While previous evaluation methods focused on the compliance of states with mandated processes, the CFSRs were intended to tie agency performance to the three broad goals of safety, permanency, and well-being (GAO, 2004). The CFSR process is administered by the U.S. Department of Health and Human Services (U.S. DHHS) Administration for Children and Families (ACF).

Since the CFSR process was initiated in 2001, two full rounds of the process have been completed. Each round of the CFSR reviews child welfare performance for each state, Puerto Rico, and the District of Columbia using the following three-phase process (GAO, 2004; CFR, 2007; U.S. DHHS, 2011):

- (1) *Statewide assessments.* Each state self-assesses its own performance on pre-defined federal outcome measures using administrative and qualitative data, and then submits the assessment to the ACF for review.
- (2) *On-site reviews.* A team of federal reviewers from ACF conducts a week-long site visit to each state to review a sample of case records and conduct interviews with agency and community stakeholders.
- (3) *Program improvement plans (PIPs).* States and ACF administrators jointly develop action plans for improving performance in areas of need identified during the review process.

These three steps complete the CFSR process, which is followed by a final two-year phase in which the PIPs are implemented and monitored to assess improvement related to the goals established in the PIP (GAO, 2004). If states do not show improvement in targeted areas addressed by the PIPs, they can be penalized financially through withholding of funds by the ACF. After the two-year PIP implementation period in each state, the CFSR process resumes with the next round of assessment (U.S. DHHS, 2011). Between 2001 and 2004, Round 1 was completed, with

Adoption Assistance and Child Welfare Act of 1980

The *Adoption Assistance and Child Welfare Act of 1980* (AACWA) was the first federal policy that codified the movement toward permanency planning in child welfare (Adler, 2001; Murray & Gesiriech, 2004; Testa, 2008). According to Testa (2008), public sentiments that led to the prioritization of permanency strategies were first articulated in the 1970s, when child welfare workers began to identify four major ways that foster care was not serving the needs of children. Testa (2008, p. 111) noted the following themes:

- (1) *Intended to be a temporary solution* for children who could not remain safely at home, foster care was being used as a permanent living situation.
- (2) Foster care placements could be changed at any time based on agency needs *at the expense of continuity for the children*.
- (3) Foster care did not offer a fundamental “*sense of belonging*” in a legally and socially sanctioned family.
- (4) The stigma of foster care did not provide children with *a respected social role*.

Based on these deficiencies of foster care, the permanency planning movement was launched with the passage of the AACWA in 1980 (Adler, 2001; Murray & Gesiriech, 2004; Testa, 2008). AACWA strongly emphasized the preservation of families by mandating that child welfare agencies make “reasonable efforts” to prevent removal of children when possible, and to expeditiously reunify families when removal could not be prevented (Berrick, 2009; Murray & Gesiriech, 2004; Testa, 2008). For children who could not be reunified with their families, AACWA also emphasized permanent placement with alternative families through adoption, although AACWA discouraged termination of parental rights except in the most extreme circumstances (Adler, 2001). Before the enactment of AACWA, there was little oversight of the child welfare and foster care systems; AACWA introduced outcome-oriented and time-limited permanency goals for children in the child welfare system and mandated state compliance with these goals (Reed & Karpilow, 2009).

Social Security Amendments of 1994, MEPA, and IEPA

The mid 1990s brought several key policy changes that would lay the groundwork for the next piece of major reform legislation enacted in 1997. The *Social Security Amendments of 1994* mandated that the Department of Health and Human Services develop regulations for the evaluation of state child welfare agencies (Corrected Federal Register, 2007, cited hereinafter as CFR, 2007), a mandate that would later result in the development of the current outcome evaluation system. The *Multi-Ethnic Placement Act of 1994* (MEPA) and the *Inter-Ethnic Placement Act of 1996* (IEPA) prohibited federally-funded agencies from considering race or ethnicity in decisions regarding the placement of children (Murray & Gesiriech, 2004). Referring to Testa’s (2008) framework, MEPA and IEPA clearly reflect a preference toward bureaucratic over primordial system organization.

Adoption and Safe Families Act of 1997

Despite the intentions set forth under AACWA in 1980, by the mid 1990s it was apparent that children were still entering and unnecessarily languishing in out-of-home care, highlighting the need to develop other permanency strategies (Adler, 2001). The federal legislative response to these continuing concerns was the passage of the *Adoption and Safe Families Act of 1997* (ASFA) that emphasized adoption as a means to achieve permanency for children in foster care (Adler, 2001; Berrick, 2009). ASFA addressed concerns that birth parents were being given unwarranted time and opportunity to reunify with their children by narrowing the time frame in which reunification

every state finishing the three phases of the CFSR process (U.S. DHHS, 2011). Between 2007 and 2010, all jurisdictions completed Round 2 (U.S. DHHS, 2011).

The first two phases of the CFSR process provide the information that the ACF uses to assess a state's performance and identify areas in which there is not conformity with established outcomes. The evaluation process is intended to measure states' performance in relation to seven outcomes using administrative data and seven systemic factors related to agency compliance with mandated procedural operations. The seven outcomes are organized under the three overarching goals of child safety, permanency, and well-being as follows (U.S. DHHS, 2011):

Safety Outcomes

1. Children are first and foremost protected from abuse and neglect.
2. Children are safely maintained in their homes when possible.

Permanency Outcomes

1. Children have permanency and stability in their living situations.
2. The continuity of family relationships and connections is preserved.

Well-Being Outcomes

1. Families have enhanced capacity to provide for children's needs.
2. Children receive services to meet their educational needs.
3. Children receive services to meet their physical and mental health needs.

Although there are seven outcomes established by the ACF, only two have related quantitative data indicators, while the other five outcomes are measured qualitatively through the case review process. The first safety outcome and the first permanency outcome are the only two outcomes that use quantitative administrative data linked to national performance standards; there are two data indicators associated with the safety outcome, and four composite data indicators associated with the permanency outcome, as noted in Figure 2 (CFR, 2007; U.S. DHHS, 2011).

In the next section, the evolution of the CFSR and the federal performance measures is examined through the lens of the conceptual framework summarized in the first section of this article and outlined in Figure 1. The focus of the analysis is on broad features of the performance measurement system, including the process by which the system was developed, rather than on technical concerns, which have been addressed extensively by Schuerman and Needell (2009) and Courtney, Needell, and Wulczyn (2004).

POLICY DEVELOPMENT AND CHILD WELFARE PRACTICE

The shift to a performance measurement system that focuses on measurable outcomes and outlines specific standards and outcomes has been viewed positively by many in the field, who have welcomed a focus on service outcomes rather than on compliance with mandated processes. (CFR, 2007; Poertner et al., 2008). However, serious concerns have also been raised in the debate at the federal level as reflected in the record submitted by U.S. DHHS-ACF to the Federal Register. Commentary from past legislative debates, beginning with the original set of regulations issued in 2000 and followed by the commentary that was submitted surrounding the issuance of regulations for Round 2 of the CFSR in 2007 provide an important context for understanding the child welfare performance measurement system. Commentary responding to the Round 1 outcomes and

Outcome	Measure
Children are first and foremost protected from harm and neglect	<p>Of all children who were victims of substantiated or indicated abuse or neglect during the first 6 months of the reporting year, what percent did not experience another incident of substantiated or indicated abuse or neglect within a 6-month period?</p> <p>Of all children in foster care during the reporting period, what percent were not victims of a substantiated or indicated maltreatment by foster parents or facility staff members?</p>
Children have permanency and stability in their living situation	<p>Permanency Composite 1: Timeliness and permanency of reunification</p> <p>C1.1 Of all children who were discharged from foster care to reunification in the target 12-month period, and who had been in foster care for 8 days or longer, what percent were reunified in less than 12 months from the date of the latest removal from home?</p> <p>C1.2 Of all children who were discharged from foster care to reunification in the 12-month target period, and who had been in foster care for 8 days or longer, what was the median length of stay in months from the date of latest removal from home until the date of discharge to reunification?</p> <p>C1.3 Of all children who entered foster care for the first time in the 6-month period just prior to the target 12-month period, and who remained in foster care for 8 days or longer, what percent were discharged from foster care to reunification in less than 12 months from the date of latest removal from home?</p> <p>C1.4 Of all children who were discharged from foster care to reunification in the 12-month period prior to the target 12-month period, what percent reentered foster care in less than 12 months from the date of discharge?</p> <p>Permanency Composite 2: Timeliness of adoptions</p> <p>C2.1 Of all children who were discharged from foster care to a finalized adoption during the 12-month target period, what percent were discharged in less than 24 months from the date of the latest removal from home?</p> <p>C2.2 Of all children who were discharged from foster care to a finalized adoption during the 12-month target period, what was the median length of stay in foster care in months from the date of latest removal from home to the date of discharge to adoption?</p> <p>C2.3 Of all children in foster care on the first day of the 12-month target period who were in foster care for 17 continuous months or longer, what percent were discharged from foster care to a finalized adoption by the last day of the 12-month target period? (The denominator for this measure excludes children who, by the last day of the 12-month target period, are discharged from foster care with a discharge reason of reunification with parents, living with other relatives, or guardianship.)</p> <p>C2.4 Of all children in foster care on the first day of the 12-month target period who were in foster care for 17 continuous months or longer, and who were not legally free for adoption prior to that day, what percent became legally free for adoption during the first six months of the 12-month target period? (The denominator for this measure excludes any child who did not become legally free during the first six months of the target year, but who, during that six-month period, is discharged from foster care with a discharge reason of reunification with parents or primary caretakers, living with other relatives, or guardianship.)</p> <p>C2.5 Of all children who became legally free for adoption during the 12 months prior to the target 12-month period, what percent were discharged from foster care to a finalized adoption in less than 12 months from the date of becoming legally free?</p> <p>Permanency Composite 3: Permanency for children in foster care for long periods</p> <p>C3.1 Of all children who were in foster care for 24 months or longer on the first day of the 12-month target period, what percent were discharged to a permanent home by the last day of the 12-month target period and prior to their 18th birthday?</p> <p>C3.2 Of all children who were discharged from foster care during the 12-month target period, and who were legally free for adoption at the time of discharge, what percent were discharged to a permanent home prior to their 18th birthday?</p> <p>C3.3 Of all who either (1) were, prior to age 18, discharged from foster care during the 12-month target period with a discharge reason of emancipation, or (2) reached their 18th birthday while in foster care but had not yet been discharged from foster care, what percent were in foster care for 3 years or longer?</p> <p>Permanency Composite 4: Placement Stability</p> <p>C4.1 Of all children who were served in foster care during the 12-month target period, and who were in foster care for at least 8 days but less than 12 months, what percent had two or fewer placement settings?</p> <p>C4.2 Of all children who were served in foster care during the 12-month target period, and who were in foster care for at least 12 months but less than 24 months, what percent had two or fewer placement settings?</p> <p>C4.3 Of all children who were served in foster care during the 12-month target period, and who were in foster care for at least 24 months, what percent had two or fewer placement settings?</p>

FIGURE 2 Round II federal child welfare performance measures (CFR, 2007).

measures included 176 letters submitted by state and local child welfare agencies, national and local advocacy groups, academic institutions, and individual social workers (Title IV–E Foster Care Eligibility Reviews, 2000, cited hereinafter as 45 CFR, 2000), as well as comments received from child welfare providers, judicial professionals, local community organizations, and subsequent focus groups with child welfare stakeholders conducted by U.S. DHHS-ACF in 2000.

Prominent themes in the discussion surrounding the multiple cycles of debate on the development of the child welfare performance measurement system, include: (1) the role of stakeholder groups in the development of performance measures; (2) the interaction and tensions between the accountability aims of the performance measurement system at the federal level, and the performance management objectives operating at the local level; and (3) the role of contextual environmental factors in the development of performance measures and the integration of those measures in child welfare practice.

Stakeholder Involvement

Evidence of balanced stakeholder involvement in the development of the CFSR and performance measures is mixed. For example, during Round 1, a number of requests were made to require that the CFSR review teams involving representatives from particular groups, including representatives of citizen review panels, representatives of the Office for Civil Rights within U.S. DHHS, or skilled child welfare practitioners. However, U.S. DHHS decided not to regulate the composition of the review team, but incorporated several provisions in the rule which they argued would ensure representativeness (45 CFR, 2000, p. 4039). For example, under 1355.33(a)(2)(ii) and (iv) states were required to list the external participants involved in the statewide process, to provide evidence that adequate consultation with stakeholders took place. In another Round 1 example, U.S. DHHS did not specify in the regulations all the individuals who should be interviewed, although they noted that case specific “[i]nterviews [would] include parents and adoptive parents [would] be routinely interviewed” (45 CFR, 2000, p. 4039). They further explained that the rule did not specify community stakeholders who should be interviewed in addition to the case specific interviews, but stated broadly that representatives with both local and statewide perspectives would be interviewed. A list of community representatives to be interviewed was provided in the separate procedures manual issued by U.S. DHHS, but this matter was not regulated (45 CFR, 2000, p. 4039).

Regarding the performance measures, U.S. DHHS noted in Round 2 that “[t]he majority of respondents to the Federal Register notice expressed support for our proposal to use data composites” (CFR, 2007, pp. 5–6). Moreover, input from stakeholders was cited as justification for decisions made by U.S. DHHS regarding particular measures. For example, in the case of the definition of placement stability, when some commentators argued for defining stability as involving three rather than two placement settings, U.S. DHHS defended the two placement or less definition of stability in part by arguing that “the existing definition of stability was established in consultation with key stakeholders in the child welfare field.” However, this account by U.S. DHHS of stakeholder input in Round 2 fails to acknowledge the substantial criticisms of the specific measures raised by other experts in the field.

Interactions between Measuring and Managing Performance

In developing the CFSR process, including defining the data based outcome measures, setting the national standards, and establishing the role these would play in the overall CFSR performance measurement system, U.S. DHHS consistently articulated its intent to balance state accountability with flexibility. U.S. DHHS also reiterated its intent to:

... maintain the flexibility to make appropriate changes that support the results-focused approach to Federal reviews of State programs ... [affording] both the Federal government and the States an ongoing opportunity to benefit from lessons learned in future reviews and make improvements to the process where needed. (45 CFR, 2000, p. 4024)

The contrast that U.S. DHHS emphasized between accountability and flexibility reflects the tension between performance measurement (accountability) and performance management (local flexibility) outlined in Figure 1. Key examples of this ongoing tension related to efforts to impose financial penalties for failure to meet standards; and mandate compliance with systemic factors in addition to outcome measures.

Addressing comments regarding the penalties to be assessed for failure to achieve substantial conformity, U.S. DHHS argued that its federal stewardship role required an adequately aggressive approach to penalizing states not in conformity, while noting that the imposition of financial penalties had consequences for states' capacity to administer child welfare programs (45 CFR, 2000, p. 4028). Ultimately, U.S. DHHS made several revisions to the penalty provisions that were intended to balance flexibility and accountability, including: (1) graduated penalties for states remaining in non-conformity after successive reviews, with maximum withholding imposed upon states that did not develop/implement a State Program Improvement Plans (PIP); (2) requirement of substantial rather than total conformity, but also penalties for each factor or outcome for which the state was not in substantial conformity; and (3) time limited opportunities for states to make improvements before penalties would be withheld from federal funding to the state (45 CFR, 2000, p. 4044) along with the imposition of penalties when the state failed to make necessary progress on specific data indicators by the time specified in the PIP (45 CFR, 2000, p. 4045).

Regarding the CFSR's establishment of systemic factors, some commentators on CFSR development suggested that if a state were in substantial conformity regarding child welfare outcomes, then it should not be held accountable for systemic factors. U.S. DHHS rejected this suggestion, responding first that the purpose of the CFSR was also to determine compliance with state plan requirements, which relate to the systemic factors. Further, U.S. DHHS noted that while procedural review was inadequate to ensure outcomes, it was essential to ensuring state capacity to deliver services most likely to achieve outcomes (45 CFR, 2000, pp. 4041–4042). This act represented an active attempt by U.S. DHHS to assert control over the performance management processes taking place within state systems. U.S. DHHS explained that it used the CFSR to help "identify areas where needed improvements [could] lead to better outcomes" (45 CFR, 2000, p. 4042). More accurately, it used the CFSR to mandate a series of practices and systems it deemed to foster improved outcomes. However, U.S. DHHS acknowledged the limitations of this endeavor, noting that "systemic changes that [led] to identifiable improvements in the outcomes for children and families [could not] always be achieved by simply modifying a policy, creating new tracking procedures, or implementing new standards" (45 CFR, 2000, p. 4043).

The intersection of federal performance measurement accountability priorities and local performance management constraints and objectives emerge repeatedly in the record of commentary on the proposed final rule for Round 2. In some instances, U.S. DHHS incorporated the state's performance management aims directly. For example, in the measures related to outcomes for children in long-term care, U.S. DHHS decided to include guardianship as one of the permanency options noting "because we recognize that many States have made concerted efforts to achieve permanency for children through guardianship, we included guardianship as a permanency option in the two measures that assess achieving permanency for children" (CFR, 2007, p. 27).

In other instances, U.S. DHHS declined to respond to critiques from local stakeholders, emphasizing instead the aims and constraints of the existing performance measurement system. For example, U.S. DHHS declined to make changes in the 12-month time frame utilized in the measure of timely reunification, although some respondents "suggested that a 12-month timeframe [was]

not sufficient in many cases to achieve reunification, particularly for families in which parental substance abuse was a key reason for a child's removal from the home" (CFR, 2007, p. 18). Although substance abuse services required the local agency to collaborate with other service systems and providers, and substance abuse problems were inherently chronic and frequently unresponsive to treatment, U.S. DHHS asserted what they termed "the responsibility of child welfare agencies to return children to safe homes as quickly as possible," and declined to extend the time frame to 18 or 24 months as requested (CFR, 2007, p. 18).

Accounting for Contextual Factors

Responses to the rules issued for Round 2 of the CFSR identified several key issues involving the relationship between environmental or contextual factors and state performance. The first issue related to the use of standards based on national data from all the states to evaluate the performance of individual states. Respondents argued that a comparative standard was not valid because "variations in State practices, statutes, and policies often impact the comparability of performance on a particular measure" (CFR, 2007, p. 9). While acknowledging the significance of variations in policies and practices among the states, U.S. DHHS took the position that they had "attempted to address these variations both in the new measures proposed for the composites and in the use of composites themselves" (CFR, 2007, p. 9). Further, they noted that the "assessment of a State's performance on its individual PIP [was], and [would] continue to be, based on change in an individual State's performance over time rather than on whether the State [met] the national standard" explaining that states were only required to make progress toward the standard in order to avoid financial penalties (CFR, 2007, p. 9). In taking this position, U.S. DHHS essentially shifted the focus of accountability to the performance management process that took place at the state level, asserting that the locally established goals and timelines were the basis for financial penalties, rather than the federally mandated outcome benchmarks.

Another issue relating to contextual variation between the states involved population and caseload characteristics. U.S. DHHS reported that "[m]any respondents to the Federal Register notice suggested that ACF should [have] assess[ed] performance on the composites and the measures to determine whether there [were] differences in performance as a result of children's age, race/ethnicity, or reasons for entering foster care and that the national standards should [have] be[en] adjusted accordingly" (CFR, 2007, p. 10). In this instance, U.S. DHHS declined to establish separate performance standards for children of different ages, races, or reasons for entering foster care, asserting that "all children [had] the same need for safety, placement stability, and timely permanency" (CFR, 2007, p. 10). Once again, U.S. DHHS emphasized local performance management processes as the appropriate way to respond to the performance issues associated with a higher prevalence of children with risk factors arguing that: "this type of analysis [was] best left to the States to further examine the characteristics of their own child welfare populations as part of their Statewide Assessment" (CFR, 2007, p. 10).

CONCLUSION

The evolving, dynamic history of child welfare policy in the United States is characterized by shifting goals and objectives. The current performance measurement system attempts to reflect and promote progress toward the system's established goals. The theoretical and empirical literature examining performance measurement and accountability in the public sector highlight the importance of: (1) ensuring stakeholder involvement, (2) balancing flexibility/local performance management priorities with accountability/external performance measurement mandates, and (3) attending to

local context when establishing performance standards. This analysis of the debate surrounding the development of the federal performance measures confirms the importance of these issues in the context of child welfare policy, but points out their inadequate incorporation in the decision making process. Future reforms should strengthen their weight in making determinations about the CFRS process and the specific performance measures utilized in the reviews.

REFERENCES

- Abma, T. A., & Noordegraaf, M. (2003). Public managers amidst ambiguity: Towards a typology of evaluative practices in public management. *Evaluation, 9*, 285–306.
- Adler, L. S. (2001). The meanings of permanence: A critical analysis of the Adoption and Safe Families Act of 1997. *Harvard Journal on Legislation, 38*, 1.
- Albers, E., Reilly, T., & Rittner, B. (1993). Children in foster care: Possible factors affecting permanency planning. *Child & Adolescent Social Work Journal, 10*, 329–341.
- Behn, R. D. (2002). Market-based governance: Supply side, demand side, upside, and downside. In J. D. Donahue & J. S. Nye, Jr. (Eds.), *Government performance and the conundrum of public trust* (pp. 323–348). Cambridge, MA: Visions of Governance in the Twenty-First Century; Washington, DC: Brookings Institution Press.
- Benjamin, L. (2008). Evaluator's role in accountability relationships: Measurement technician, capacity builder or risk manager? *Evaluation, 14*, 323–343.
- Berrick, J. D. (2009). *Take me home: Protecting America's vulnerable children and families*. New York, NY: Oxford University Press.
- Bovens, M., Schillemans, T., & Hart, P. (2008). Does public accountability work? An assessment tool. *Public Administration, 86*, 225–242.
- Broad, M., Goddard, A., & Von Alberti, L. (2007). Performance, strategy and accounting in local government and higher education in the UK. *Public Money & Management, 27*, 119–126.
- Corrected Federal Register Announcement [CFR]. (2007). United States Department of Health and Human Services. Retrieved from http://www.acf.hhs.gov/programs/cb/cwmonitoring/legislation/fed_reg.pdf
- 45 CFR. (2000). Title IV–E Foster Care Eligibility Reviews and Child and Family Services State Plan Reviews; Final Rule. Department of Health and Human Services—Administration for Children and Families, 45 CFR Parts 1355, 1356 and 1357. Retrieved from http://www.acf.hhs.gov/programs/cb/laws_policies/cblaws/fed_reg/fr012500.pdf
- Courtney, M., Needell, B., & Wulczyn, F. (2004). Unintended consequences of the push for accountability: The case of national child welfare performance standards. *Children and Youth Services Review, 26*, 1141–1154.
- Dahler-Larsen, P. (2007). Constitutive effects of performance indicator systems. In S. Kushner & R. Stake (Eds.), *Dilemmas of engagement: Evaluation and the new public management (Advances in program evaluation, Volume 10)* (pp. 17–35). Bingley, UK: Emerald Group Publishing Limited.
- Ebrahim, A. (2005). Accountability myopia: Losing sight of organizational learning. *Nonprofit and Voluntary Sector Quarterly, 34*, 56–87.
- Fellin, P. (2001). *The community and the social worker* (3rd ed.). Independence, KY: Davis: Brooks/Cole, Cengage Learning.
- Forbes, D. P. (1998). Measuring the unmeasurable: Empirical studies of nonprofit organization effectiveness from 1977 to 1997. *Nonprofit and Voluntary Sector Quarterly, 27*, 183–202.
- Fox, A., Frasch, K., & Berrick, J. (2000). *Listening to children in foster care: An empirically based curriculum*. UC Berkeley: Child Welfare Research Center. Retrieved from http://cssr.berkeley.edu/research_units/cwrc/publications_details.html
- General Accounting Office. (2004). *Child and Family Services Reviews: Better use of data and improved guidance could enhance HHS's oversight of state performance*. (GAO 04-333). Washington, DC: Author.
- Hasenfeld, Y. (1983). *Human service organizations*. Upper Saddle River, NJ: Prentice-Hall, Inc.
- Hatry, H., van Houten, T., Plantz, M., & Greenway, M. (1996). *Measuring program outcomes: A practical approach*. Alexandria, VA: United Way of America.
- Hatry, H. P. (2002). Performance measurement: Fashions and fallacies. *Public Performance & Management Review, 25*, 352–358.
- Hatry, H. P. (2006). *Performance measurement: Getting results* (2nd. ed.). Washington, DC: The Urban Institute Press.
- Heinrich, C. (2002). Outcomes-based performance management in the public sector: Implications for government accountability and effectiveness. *Public Administration Review, 62*, 712–725.

- Heinrich, C. (2007). Evidence-based policy and performance management. *American Review of Public Administration*, 37, 255–277.
- Kaplan, R. S. (2001). Strategic performance measurement and management in nonprofit organizations. *Nonprofit Management and Leadership*, 11, 353–370. doi: 10.1002/nml.11308
- Kemp, S. P., & Bodonyi, J. M. (2002). Beyond termination: Length of stay and predictors of permanency for legally free children. *Child Welfare*, 81, 58–86.
- Melkers, J., & Willoughby, K. (2005). Models of performance-measurement use in local governments: Understanding budgeting, communication, and lasting effects. *Public Administration Review*, 65, 180–190. doi: 10.1111/j.1540-6210.2005.00443.x
- Moxham, C., & Boaden, R. (2007). The impact of performance measurement in the voluntary sector: Identification of contextual and processual factors. *International Journal of Operations and Production Management*, 27(8), 826–845.
- Moynihan, D. P., & Pandey, S. K. (2005). Testing how management matters in an era of government by performance management. *Journal of Public Administration Research and Theory*, 15, 421–439.
- Murray, K., & Gesirich, S. (2004). *A brief legislative history of the child welfare system*. Washington, DC: Pew Charitable Trusts. Retrieved from http://www.pewtrusts.org/our_work_report_detail.aspx?id=48994
- Needell, B., & Putnam-Hornstein, E. (2009a). Understanding data and the child welfare system. *Insights*, 1. Retrieved from http://www.co-invest.org/resources/Insights_0809.pdf
- Needell, B., & Putnam-Hornstein, E. (2009b). Understanding permanency and the child welfare system. *Insights*, 2. Retrieved from http://www.co-invest.org/resources/Insights_Vol2_Dec09.pdf
- Park, J. M., & Ryan, J. P. (2009). Placement and permanency outcomes for children in out-of-home care by prior inpatient mental health treatment. *Research on Social Work Practice*, 19, 42–51 (originally published online April 29, 2008). doi: 10.1177/1049731508317276
- Perrin, B. (1998). Effective use and misuse of performance measurement. *American Journal of Evaluation*, 19, 367–379.
- Pecora, P. (2006). Child welfare policies and programs. In J. M. Jensen & M. W. Fraser (Eds.), *Social policy for children and families* (pp. 19–66). New York, NY: Oxford University Press.
- Poertner, J., Moore, T., & McDonald, T. P. (2008). Managing for outcomes: The selection of sets of outcome measures. *Administration in Social Work*, 32, 5–22.
- Reed, D., & Karpilow, K. (2009). *Understanding the child welfare system in California: A primer for service providers and policymakers* (2nd ed.). Berkeley, CA: California Center for Research on Women and Families. Retrieved from <http://www.ccrwf.org/2009/07/understanding-the-child-welfare-system-in-california-a-primer-for-service-providers-and-policymakers/>
- Salaman, G., & Storey, J. (2002). Managers' theories about the process of innovation. *Journal of Management Studies*, 39, 147–165. doi: 10.1111/1467-6486.00286.
- Sanger, M. B. (2008). From measurement to management: Breaking through the barriers to state and local performance. *Public Administration Review*, 68, S70–S85. doi: 10.1111/j.1540-6210.2008.00980.x
- Schalock, R. L., & Bonham, G. S. (2003). Measuring outcomes and managing for results. *Evaluation and Program Planning*, 26, 229–235.
- Schuerman, J., & Needell, B. (2009). *The child and family services review composite scores: Accountability off the track*. Chicago, IL: Chapin Hall at the University of Chicago.
- Speckbacher, G. (2003). The economics of performance management in nonprofit organizations. *Nonprofit Management and Leadership*, 13, 267–281. doi: 10.1002/nml.15
- Testa, M. (2008). New permanency strategies for children in foster care. In D. Lindsey & A. Shlonsky (Eds.), *Child welfare research* (pp. 108–124). New York, NY: Oxford University Press.
- U.S. Department of Health and Human Services. (2011). Child and Family Services reviews aggregate report: Findings for round 2, fiscal years 2007–2010. Retrieved from http://www.acf.hhs.gov/sites/default/files/cb/fcfsr_report.pdf
- Van Thiel, S., & Leeuw, F. L. (2002). The performance paradox in the public sector. *Public Performance & Management Review*, 25, 267–281.
- Ward, D. E., Hamilton, J., Fein, E., & Maluccio, A. (1982). Planning for permanency planning in foster care. *Children and Youth Services Review*, 4, 223–237.
- Wells, S. J., & Johnson, M. A. (2001). Selecting outcome measures for child welfare settings: Lessons for use in performance management. *Children and Youth Services Review*, 23, 169–199.
- Woods, M., & Grubnic, S. (2008). Linking comprehensive performance assessment to the balanced scorecard: Evidence from Hertfordshire County council. *Financial Accountability & Management*, 24, 343–361.
- Yetano, A. (2009). Managing performance at local government level: The cases of the City of Brisbane and the City of Melbourne. *Australian Journal of Public Administration*, 68, 167–181. doi: 10.1111/j.1467-8500.2009.00632.x.