

2008

# Legal Accountability in the Service-Based Welfare State: Lessons from Child Welfare Reform

Kathleen G. Noonan

*University of Pennsylvania*, noonan@sp2.upenn.edu

Charles F. Sabel

*Columbia Law School*, csabel@law.columbia.edu

William H. Simon

*Columbia Law School*, wsimon@law.columbia.edu

Follow this and additional works at: [https://scholarship.law.columbia.edu/faculty\\_scholarship](https://scholarship.law.columbia.edu/faculty_scholarship)



Part of the [Family Law Commons](#), [Public Law and Legal Theory Commons](#), and the [Rule of Law Commons](#)

---

## Recommended Citation

Kathleen G. Noonan, Charles F. Sabel & William H. Simon, *Legal Accountability in the Service-Based Welfare State: Lessons from Child Welfare Reform*, *LAW OF SOCIAL INQUIRY*, VOL. 34, P. 523, 2009; COLUMBIA PUBLIC LAW RESEARCH PAPER NO. 08-162 (2008).

Available at: [https://scholarship.law.columbia.edu/faculty\\_scholarship/1510](https://scholarship.law.columbia.edu/faculty_scholarship/1510)

This Working Paper is brought to you for free and open access by the Faculty Publications at Scholarship Archive. It has been accepted for inclusion in Faculty Scholarship by an authorized administrator of Scholarship Archive. For more information, please contact [donnely@law.columbia.edu](mailto:donnely@law.columbia.edu).

# Legal Accountability in the Service-Based Welfare State: Lessons from Child Welfare Reform

Kathleen G. Noonan, Charles F. Sabel, and William H. Simon

*Current trends intensify the longstanding problem of how the rule of law should be institutionalized in the welfare state. Welfare programs are being redesigned to increase their capacities to adapt to rapidly changing conditions and to tailor their responses to diverse clienteles. These developments challenge the understanding of legal accountability developed in the Warren Court era. This article reports on an emerging model of accountable administration that strives to reconcile programmatic flexibility with rule-of-law values. The model has been developed in the reform of state child protective services systems, but it has potentially broad application to public law. It also has novel implications for such basic rule-of-law issues as the choice between rules and standards, the relation of bureaucratic and judicial control, the proper scope of judicial intervention into dysfunctional public agencies, and the justiciability of “positive” (or social and economic) rights.*

## INTRODUCTION

Recent trends in welfare systems in America and abroad intensify longstanding uncertainties about how rule-of-law values apply in these systems.

---

**Kathleen Noonan** is Managing Director, PolicyLab, Children’s Hospital of Philadelphia. Contact: noonank@email.chop.edu.

**Charles F. Sabel** is Maurice T. Moore Professor of Law, Columbia University. Contact: cfs11@columbia.edu.

**William H. Simon** is Arthur Levitt Professor of Law, Columbia University. Contact: wsimon@law.columbia.edu.

The authors are grateful to many people for advice and encouragement, including Richard Anderson, Stephen Cohen, Navina Forsythe, Barbara Fried, Ivor Groves, Jonathan Lipson, Jerry Mashaw, Brad McGarry, Sherry Orbach, Carol Spigner, Jane Spinak, George Taylor, Paul Vincent, and Michael Wald.

Programs that once focused on financial redistribution increasingly link transfer payments to services, and services are increasingly customized to the needs of individual recipients. The move to services is driven by the perception that transfer payments alone do not induce (and may inhibit) the development of skills that permit self-sufficiency. The move to individuation is driven in part by a conception of fairness that mandates response to “difference” in people’s values and circumstances and by the perception that these circumstances are more fluid than they have been in the past (Handler 2004).

As the welfare state becomes more individuating and more adaptive, it threatens to undermine the awkward compromise between rule-of-law values and welfare state practices worked out in the Warren Court years and their aftermath. Two key features of that compromise were (1) the idea of a balance between relatively rigid rules to govern the conduct of low-status frontline workers and relatively flexible standards to govern the conduct of professionals, and (2) the idea of coordination between a bureaucratic accountability system for routine cases and a quasi-judicial accountability system for cases in which beneficiaries protest their treatment (Simon 1984). In addition, the compromise distinguished two modes of court intervention into the administrative system—routine discrete intervention focused on particular practices or narrow norms and extraordinary systemic intervention designed to restructure entire programs (Chayes 1976).

The core tendencies of the new programs put these arrangements under pressure. The need to customize and adapt makes rules an ineffective means of controlling discretion. Effective review of frontline efforts routinely requires the type of beneficiary participation that the old regime reserved for cases in which beneficiaries complained. Because the emerging system involves more complex coordination and more frequent adjustment, judicial review of discrete judgments and practices seems less practicable.

Yet, at the same time they create new pressures, current developments suggest new opportunities. In this article, we explore the possibility of a novel and promising accommodation of rule-of-law values and the new welfare state. We focus on developments in child protective services, especially in Alabama and Utah. Child protective services may seem an unlikely realm in which to discover rule-of-law success. Doctrinally, the field is dominated by vague standards such as “substantial risk of harm” that connote uncabinable discretion. Institutionally, the field has been associated with chaos, oppression, and tragic ineffectiveness. As we elaborate in Part I, a series of major federal statutory initiatives failed to impose order on the state-run systems. In at least thirty states, courts have found or defendants have conceded systemic noncompliance with constitutional or statutory requirements on a scale warranting structural intervention (ABA Center on Children and the Law 2005, 1). However, some of these interventions have made progress, and the model we find promising has emerged in a handful of states. It may be because these systems have been so deeply broken that they have lent themselves to

relatively radical experimentation. Or perhaps because child welfare has always been committed in principle to the individuation and adaptability that has only recently characterized the welfare state generally, it has proven fertile ground for innovation that combines these qualities with accountability.

Our aims are twofold. First, in Parts II and III, we report on the consolidation of a distinctive model of child welfare administration and suggest that it promises improved performance (though we have only impressionistic evidence for this promise). Some elements of this model have been themes in child welfare discussion for many years, but one represents an important innovation—a process of diagnostic monitoring called the Quality Service Review (QSR). The QSR is an important contribution to the approach to public administration we call “experimentalist”—an approach that seeks to induce continuous reconsideration of a system’s norms in the course of monitoring compliance with them (Dorf and Sabel 1998; Sabel 1995; Simon 2006).

Second, especially in Part IV, we draw attention to the jurisprudential properties of this model and argue that they complicate the longstanding debates about the rule of law in the welfare state and suggest possible resolutions of those debates. The Alabama-Utah model suggests a response to the question of the optimal specificity of legal norms—rules versus standards—that combines the accountability associated with rules with the contextualization associated with standards. It is also pertinent to the choice between bureaucratic and adjudicatory modes of administrative control. The central process in this model combines features of both bureaucracy and adjudication in ways that have been occasionally called for but rarely observed in the literature on legal accountability in the welfare system (see Handler 1986).

Furthermore, Alabama and Utah have implications for the debate over the proper scope of judicial intervention into chronically underperforming public institutions (Chayes 1976; Sabel and Simon 2004). Reform in these states emerged from judicial decrees mandating broad institutional form; yet, in each case the court and the parties avoided the rigidification and arbitrariness associated with “command-and-control” type judicial intervention.

Finally, the reforms we discuss have implications for the debate over the nature of welfare rights. Legal tradition makes a basic distinction between “negative rights” to be free from state interference and “positive” rights to state assistance. Theory is often torn between, on the one hand, the insight that any strong version of this distinction seems arbitrary in the light of the relative importance of the social interests that a modern legal system should protect and, on the other hand, recognition that the traditional notions of right do not seem fully generalizable to the welfare system. But the reforms we describe resonate with a conception of legal right that is responsive to the interests created by the modern welfare state and capable of effective institutionalization across the legal system. The conception has been observed in

the jurisprudence of the South African Constitutional Court (Tushnet 2003), which sees welfare rights as connoting, most fundamentally, entitlement not to a particular outcome or benefit but to a process in which the relation between the claimant's interests and the values underpinning the relevant public programs can be fairly and effectively considered. The Alabama-Utah model is richly suggestive as to how this notion of rights, only vaguely invoked in the celebrated South African cases, might be elaborated.

## I. THE STRUGGLE FOR CHILD PROTECTION

### The Federal Response to Abuse and Neglect

Abused and neglected children were identified as a social problem at the end of the nineteenth century by lay philanthropies, most notably a group of local Societies for the Prevention of Cruelty to Children (SPCC), the best known of which were in New York and Boston. They gradually evolved two competing perspectives. The "rescue" perspective, associated with the New York SPCC, emphasized intervention and removal of children from homes with abusive or neglectful parents, typically to institutions. The "preventive" approach, associated with the Boston SPCC, emphasized in-home support through social services and material assistance (Costin, Karger, and Stoesz 1996).

The public assistance titles of the Social Security Act of 1935 created a federal program of grants-in-aid to the states to support income maintenance and social services for "dependent children." Foster care, one of the federally supported services available through welfare workers, evolved into a routine response to severe family problems, including abuse and neglect.

Concern about child abuse and neglect intensified in the 1960s and 1970s when the term "battered-child syndrome" was introduced into medical diagnosis. The federal Child Abuse Prevention and Treatment Act of 1974 created a federal system for collecting data on child abuse, promulgated a model state reporting statute, and directed federal money to state programs protecting at-risk children. Reported and documented instances of abuse soared. So did foster care placements.

A "preservationist" reaction soon emerged, as did a general impression that the system was out of control. Critics complained that children were arbitrarily and unnecessarily removed to foster care. Once there, they might be shifted repeatedly from placement to placement, or simply left alone without monitoring and reassessment. Investigations showed shocking administrative disarray in the state systems; many simply could not account at all for large numbers of children that they had taken charge of. And even when child welfare agencies were minimally accountable, the routinized nature of their responses was cause for concern; in practice, workers used a

very small menu of interventions that took little account of particular circumstances (Garrison 1987).

In 1980, Congress overhauled the child protection regime with the Adoption Assistance and Child Welfare Act (AACWA), which set conditions for federal grants for child welfare services, including foster care, adoption, and family support. The act, as frequently amended, continues to provide the basic federal framework.

Substantively, AACWA declares “permanency” as the predominant goal for children in state care. It also mandates priority for natural family preservation and, where that is not possible, for adoption rather than foster care, and for any kind of family care rather than institutional care.

Procedurally, AACWA, since 1980, has prescribed the kind of individuated attention that has become a central goal of most welfare programs. The caseworker, for example, must prepare a “case plan” for each foster care placement that explains how the permanency and other goals of the statute are being met. The plan draws on an array of services; the record must document the “appropriateness” of the services provided, and the child’s circumstances and the plan must be reviewed “periodically but no less frequently than every six months” (AACWA, 42 U.S.C. sections 671(a)(16), 675(1)).

Case work is conceived as a process of “coordination” and “collaboration” among stakeholders (parents, caregivers, and children), professionals, and institutions. States are encouraged or mandated to create “multidisciplinary teams”; to “collaborate . . . with families”; to enhance the ability of “community-based programs to integrate shared leadership strategies with parents and professionals”; to “enhance interagency collaboration between the child protection system and the juvenile justice system”; to support “collaboration among public health agencies, the child protection system, and private community-based programs”; and to foster “cooperation of State law enforcement officials, courts of competent jurisdiction, and appropriate State agencies providing human services” in responding to abuse and neglect (42 U.S.C. sections 5102a(a), 5106(b)(2)(xi)).

The dialectic of “rules” (rigid and specific norms) and “standards” (flexible and vague ones) has been a prominent theme in the AACWA. The 1980 text embraced family preservation through a standard—states were required to make undefined “reasonable efforts” to avoid removing a child from his natural parents, and where removal was necessary, to reunify the family (42 U.S.C. section 671(a)(10)). In 1997, Congress retreated somewhat from preservation. The use of foster care, which initially declined after 1980, was trending upward again. An important influence was the increase in serious abuse and neglect cases associated with the epidemic of crack cocaine. Another was the reappraisal of the “preservation” model, in part as a result of research failing to confirm its presumption that children would usually fare better if kept in their own families with supportive services than if moved to foster care (Garrison 2005).

Apparently fearing that the “reasonable efforts” standard was being treated as a virtually irrefutable presumption, Congress, in the Adoption and Safe Families Act of 1997 (ASFA), made an explicit exception to the duty to pursue reunification for situations where such efforts would impede permanency or where there were “aggravated circumstances” such as violence or sexual abuse. In order to facilitate adoption, ASFA also mandated that states initiate termination of parental rights whenever a child has been in foster care for fifteen of the most recent twenty-two months (AACWA, 42 U.S.C. section 675(5)(e)). Thus, Congress initially prescribed a standard, then perceived that it was being treated like a rule strongly favoring reunification, and so pulled back by prescribing an explicit rule limiting reunification efforts. Not surprisingly, the ASFA twenty-two-month rule is now criticized as pointlessly rigidifying the process, especially for the large number of long-term foster care children who have minimal prospect of adoption (Ross 2003; Stack 2005).

The dialectic of bureaucratic and adjudicatory control is also prominent. The original legislation emphasized judicial modes of accountability. Congress mandated as early as 1974 that children have a representative (in practice, often a lawyer) in any removal proceeding (AACWA, 42 U.S.C. section 5106a(b)(2)(13)). In 1980, AACWA required that the states provide a “fair hearing” procedure for parents or caregivers who claimed they had been improperly refused services under the act (42 U.S.C. section 671(a)(12)). Most importantly, AACWA sought to make the juvenile or family court a key monitor of administrative compliance. The appropriateness of the “case plan” must be judicially reviewed at least once every six months. There must be a judicial “permanency” hearing within twelve months of removal of a child and at least annually thereafter. And removal of a child requires a judicial determination that the “reasonable efforts” at reunification required by the statute have been made (42 U.S.C. sections 672a(2)(A)(ii); 675(5)(B), 675(5)(E)(iii)).

However, expectations for quasi-judicial and judicial oversight have been disappointed. Representatives for the child can play important roles in situations of high-stakes disputes, such as contested terminations of parental rights. But they typically have high caseloads—100 to 150—and therefore cannot often play an active role in routine decision making. “Fair hearing” systems are not a significant influence in most states. Beneficiaries are not aware of them, the systems have few resources, and they have little influence beyond a small number of cases. Most cases come before the courts, but judges typically lack the knowledge, the resources, or the inclination to undertake searching review. Despite the AACWA requirement that judges must determine that “reasonable efforts” are being made in any permanency hearing, a 1989 study found that no such determination was made in 44 percent of the cases (Guggenheim 2005, 189; Huntington 2006).

Moreover, judges who take their oversight responsibilities seriously feel constrained by the limits of case-by-case intervention. They can order

additional analysis, reject proposed placements, and mandate services, but the efficacy of these alternatives depends on the larger system. Where workers are overwhelmed, available placements tend to be unsatisfactory, and service options are narrow, judges may accept as “reasonable” efforts that would not be reasonable in a more adequate system.

In the 1990s, Congress turned to bureaucratic control. Statutes had long obliged the states to report a broad range of data on their child welfare activities, and the Administration for Children and Families of the Department of Health and Human Services had long been charged with monitoring state compliance with federal statutory requirements. Congress now sought to augment and reorient such conventional monitoring in two ways. First, the state (if it accepts certain federal support) must establish “citizen review panels” (AACWA, 42 U.S.C. section 5106(a)). The panels, composed of volunteers, including some with expertise in child welfare, are to review the overall performance of the child protection agencies in the light of the statutory goals and make an annual public report. The state must also arrange for an “independently conducted audit of its programs at least every three years” (42 U.S.C. section 671(a)(13)).

Second, and most importantly, Congress mandated what has come to be known as Child and Family Service Review (CFSR). Reflecting growing dissatisfaction with “command-and-control” regulation, amendments sought to move federal oversight from a “compliance” orientation, in which success is measured by conformity to rule toward a “performance” orientation, in which the focus is on achievement of goals. In this spirit, the statute directs the Secretary of Health and Human Services (HHS) to develop a system of federal review based on “outcome measures” that “rate” state performances and examines “the reasons for high performance and low performance” (42 U.S.C. section 679b).<sup>1</sup>

Yet, for the most part, these initiatives have not borne fruit. “Citizen review” panels have been formed, but they meet only erratically and often lack expertise and access to information. HHS’s initial efforts at outcome-oriented review were fumbling and often arbitrary (National Coalition for Child Protection Reform 2003). There are some indications that they have recently improved but only after they were reconceived along the lines pioneered in Alabama and Utah that we discuss below.

---

1. The change in orientation resonated with the then-prominent “New Public Management” literature, which aimed to improve public administration by clearly separating goal setting from implementation and using reporting and incentive schemes to induce administrative agents to achieve the ends of their political principals (see Harlow 1999). Executive Order 12,866 (3 C.F.R. 638 (1999)) required that all federal agencies, when feasible, regulate by prescribing objectives rather than specific behavior.



In general, the federal legislative mandates have failed to coherently reshape practice in the states. The progress that we report in Alabama and Utah has resulted from state-level experimentation only tenuously connected to the federal statutes.

Moreover, the debate between the “rescue” and “preservation” perspectives continues. Commentators still disagree about the weight to attach to natural family ties when parenting has been gravely dysfunctional. And they continue to disagree over whether administrative dysfunction creates greater risks of inappropriate removal (Guggenheim 2005) or of inappropriate failure to remove (Bartholet 1999).

### Public Law Litigation: False Starts

We have noted that in about two-thirds of the states, all or part of the child welfare system has been successfully challenged in lawsuits seeking systemic injunctive relief. The challenges involve demonstrations or concessions of massive noncompliance with federal requirements—failure to take action in response to indications of abuse and neglect; arbitrary removal of children without reasonable reunification efforts; and placement of children in inappropriate, often dangerous, settings without substantial consideration or review (ABA Center on Children and the Law 2005).

These lawsuits have been based on the Fourteenth Amendment due process clause,<sup>2</sup> on the AACWA,<sup>3</sup> on the federal Rehabilitation Act prohibition of discrimination in federally supported programs against the disabled, and on state laws. Regardless of the particular substantive formulations of the claims, the remedies tended to assume common patterns. For many years, the general tenor of the decrees was to restrict discretion and force action through rigid rules. In recent years, there has been a move toward standards, but both rule-based and standards-based decrees have typically encountered problems.

Chris Hanson (1994, 230) noted in 1994 that “the common wisdom among litigators is that the decree should be as specific as possible.” Marcia Lowry, who brought several landmark structural reform cases, has asserted the underlying view with exceptional bluntness:

---

2. In particular, on *Youngberg v. Romeo* (1982), which held that a developmentally disabled person in state custody had a due process claim where the conditions of confinement represented “such a substantial departure from accepted professional judgment, standards, or practice to demonstrate that the person responsible actually did not base the decision on such a judgment” (322).

3. Some were initiated prior to the Supreme Court decision in *Suter v. Artist M.* (1992), rejecting private enforcement of the statute. Some have been brought since 1994, when Congress partially overruled *Suter*. See *Charlie H. v. Whitman* 2000 (discussing the extent to which the 1994 amendment overrules *Suter*).

People who run child welfare systems cannot be left to their own devices. They will not use reasonable standards, they do have to be told “first, you put your left foot in front of your right foot, then you put your right foot in front of your left foot, then you do it again.” (Lowry 1999–2000, 453)

In this spirit, the most striking tendency in the early decrees, still prominent in some more recent ones, is a preoccupation with deadlines, quantitative measures, and specific procedural and documentation requirements. An example is the 1983 decree (later fundamentally revised) in the Missouri case of *G. L. v. Zumwalt* (1983), where Lowry represented the plaintiffs. The decree specified the amount of time that must be spent training foster care licensing workers (two days initially, one day per year thereafter), foster parents (twelve hours initially, ten hours per year thereafter in general matters; thirty hours initially in prevention of abuse and neglect), case-workers (one week initially, four days per year thereafter in general matters; sixteen hours initially in prevention of abuse and neglect), and supervisors (one week initially, two days per year thereafter). There are comparable fixed deadlines for removal decisions, visits to foster homes, and the provision of various health services; fixed maxima for the number of children in a foster home or the number of cases per worker; and detailed rules about what information must be in the file.

More recent decrees have moved away from the “command-and-control” tenor of *G. L.* toward an emphasis on performance standards in the manner urged by the New Public Management literature. The 2004 plan in New Jersey is an example. The plan includes a discussion of broad goals and principles and proclaims that it should not be understood as “a checklist of requirements to be complied with, divorced from its larger purpose and context” (New Jersey Department of Children and Families 2004a, 2004b, 2). Instead of dictating practice, it mandates the formulation of plans to accomplish goals under the supervision of a monitoring panel, and it emphasizes performance measurement as a mode of assessing compliance.

Nevertheless, these performance-oriented regimes are sometimes experienced as just as restrictive as the rule-oriented ones. The New Jersey regime includes 240 “enforceable elements” prescribing plans to be formulated and implemented within specified time frames from 6 months to 2 years. It sets forth 22 basic systemic indicators (from “length of stay for all children by entry cohort” (New Jersey Department of Children and Families 2004a, 207) to “the percentage of children who have a substantiated allegation of abuse or neglect within twelve months of exit from out-of-home care to reunification” (210) and 98 “benchmarks” (from percentage of reports of child abuse or neglect as to which the child has been interviewed within 24 hours (212) to percentage of “new supervisors [who] receive the requisite pre-service training before carrying a caseload” (234) on which specified performance levels were mandated.

The experience under these decrees has been frustrating and disappointing. Eleven years after negotiating the *G. L.* decree, the parties returned to court with a substantially reoriented regime (*G. L. v. Stangler* 1994).<sup>4</sup> The need for redirection was recognized more quickly in New Jersey. The first annual report of the monitoring panel concluded that the state had “failed to meet the commitment for this monitoring period” and improvement would require “significant course correction rather than minor adjustments” (New Jersey Department of Children and Families 2005, 12).<sup>5</sup>

The increasingly recognized difficulties with both the traditional command-and-control and the New Public Management remedial orientations are these:

(1) There is the danger of preoccupation with measurable and specifiable norms at the expense of amorphous but nonetheless more important norms. Even the most restrictive decree must fall back on initially unspecified standards for some purposes. The *G. L.* decree, in addition to mandating a medical examination within twenty-four hours of custody, mandates that the child’s health needs be “adequately met” (*G. L. v. Zumwalt* 1983, 1037). Since it is easy to determine compliance with bright-line norms, enforcers may be tempted to focus on them, while paying less attention to areas where compliance assessment will be more difficult and controversial, for example, the adequacy of the medical care. After several years under a decree resembling *G. L.*, a Utah monitoring panel complained that the agency “placed a much greater emphasis on paper work compliance and compliance with the prescriptive items of the settlement agreement than on the quality of the day-to-day work with children and families” (Utah Division of Child and Family Services 1999, 8).

The focus on specification and measurement probably accounts for the tendency of administrators and frontline workers to perceive consent decree requirements as a distraction from their core mission. Social work training emphasizes general values of child welfare, informal and qualitative judgment, and personal interaction with peers and clients. (Since several decrees encourage or mandate the hiring of workers with social work degrees or in-service social work training for current workers, they intensify both sides of the cultural conflict between command-and-control and professional autonomy.)

(2) Highly specified decrees freeze practice and inhibit adjustment in the light of experience. Modifications of a decree require the consent of the

---

4. The decree was amended again in 2001 and terminated upon a finding of substantial compliance in 2006. In general, the changes reflected the trends that were more fully developed in Alabama and Utah, which we describe below.

5. In July 2006, the plaintiffs and Governor Corzine’s administration thoroughly renegotiated the decree and initiated reforms at least partially in the spirit of those we describe below in Alabama and Utah. More recent monitoring reports have found substantial improvement (New Jersey Department of Families and Children 2008).

plaintiff or a court order, which normally requires a showing of changed circumstances (*Rufo v. Inmates of Suffolk County Jail* 1992). Such a process does not contemplate routine modification, but social service practice needs to adjust continually. Negotiation over discrete issues, even with cooperative plaintiff's counsel, can be cumbersome. The original Utah decree, for example, required that all children undergo mental health assessments. It soon became apparent that this was pointless for infants, but modification took considerable time (8).

(3) Some decrees use indicators that are useless in setting priorities in the inevitable situation when there is noncompliance. This is potentially as true as for performance-based decrees, such as New Jersey's, as for rule-based ones like *G. L.* New Jersey's first monitoring report showed poor compliance or noncompliance with the majority of its 255 "enforceable elements." But a monitoring system that reports failure from all directions cannot direct efforts to improve. Thus, the monitoring panel recommended that the parties specify and focus on a small number of core goals. The state, it said, should "attempt to do a smaller number of fundamental things and to do them very well, rather than continuing to attempt to implement all portions of the reform plan with equal priority" (New Jersey Department of Children and Families 2005, 12).

(4) An effective monitoring regime needs both outcome information that indicates how effective the system is in achieving its goals and diagnostic information that indicates the locales and practices that are responsible for failures. Much of the information produced by command-and-control decrees has little value for either purpose. For example, the number or fraction of cases in which, within a given time frame, meetings were not held or forms not completed and filed has at best indirect value as either a measure of outcomes or a locator of failure (especially if the data do not distinguish a total failure to meet or file from an inadequate meeting or filing, and are silent as to the specific nature and cause of the failure or inadequacy). New Jersey's performance-based decree emphasizes important measures of outcome—for example, frequency of reabuse in state custody, the length of time to permanency, and the frequency of "reentry" into the system after cases are closed. But by themselves, these measures have limited diagnostic value. Low scores suggest that the system is doing badly but give only the vaguest indication as to how it could improve.

## II. THE EMERGING MODEL: ALABAMA AND UTAH

A model of child welfare reform that avoids the pitfalls of command-and-control, on the one hand, and single-minded outcome-focus, on the other, is emerging.

The model has been developed with exceptional self-consciousness and efficacy in Alabama and Utah. Accordingly, we focus on developments in these states and refer to “the Alabama-Utah model.” Most features of the model can be found in other systems, and most follow longstanding recommendations in the policy literature and, in some instances, prescriptions of federal legislation. Their success in these states seems due in part to a convergence of efforts by advocates, state officials, and federal judges who eventually came to take seriously the decentralizing themes of the literature and legislation.

While much of the model represents thoughtful, disciplined implementation of established views, two features of practice in Alabama and Utah strike us as innovative. First, reformers in these two states have developed a distinctive form of diagnostic monitoring they call the Quality Service Review (QSR) that furthers the aspiration to combine contextual decision making with systemic accountability.<sup>6</sup> Second, lawyers and judges in the Alabama and Utah court cases have contributed to the elaboration of a conception of social welfare right that emphasizes a system’s capacity for self-assessment and self-correction over compliance with judicially derived substantive standards.

## Background

The landmark case is *R. C. v. Walley*, a class action challenge to the Alabama system filed in 1988 and settled in 1992. Paul Vincent, the director of the state’s division of family and child services at the time, was a progressive social worker who welcomed the lawsuit as an opportunity to rethink and restructure the system. The lead plaintiff’s counsel, Ira Burnim, of the Bazelon Center for Mental Health Law, had been involved in *Wyatt v. Stickney*, a pioneering structural reform case against Alabama’s institutions for the mentally disabled (Bazelon Center for Mental Health Law 1998). Top-down rule-based intervention did not sit well with the conception of practice Vincent had learned in social work, and Burnim had become sensitive to the limits of such approaches from his experience in *Wyatt* and other cases. They and several collaborators decided to attempt a remedial regime that would achieve accountability while accommodating decentralization and flexibility. They were later joined by Ivor Groves, a former Florida mental health administrator, who became the court-appointed monitor in *R. C.* (Bazelon Center for Mental Health Law 1998).

Vincent and Groves both contributed to efforts to redirect the reform process instigated in Utah through *David C. v. Leavitt*, which was filed in 1992

---

6. The process is sometimes referred to as “Qualitative Service Review,” or, in Utah, “Qualitative Case Review.”

and first settled in 1994. After four years under the initial settlement agreement, the parties had returned to court in frustration. Despite an increase of more than 50 percent in both state funding and the number of caseworkers, there was virtually no indication of progress on most of the measures stipulated in the agreement. In an account that foreshadowed the situation in New Jersey some years later, the court found, “the feedback mechanisms designed by the parties have proven unworkable or unrealistic” (*David C. v. Leavitt* 1998, 1212). There had been a “breakdown in the Corrective Action Process” (1207). At this point, another visionary social worker—Richard Anderson, then deputy director and later director of the state Division of Child and Family Services—appeared and worked with plaintiffs’ counsel to reorient the process. Vincent and Groves, who by now each headed private consulting organizations, were retained to monitor and assist.

### General Themes

The shift reflected in the Alabama decree and the second Utah decree has been described as “moving from a law-based practice to a social-work based practice” (Utah Division of Child and Family Services 1999, 8).<sup>7</sup> The phrase is apt in noting the influence of social work tradition, but it is misleading in two senses. First, in connoting a notion of law as rule-bound and hierarchical, it ignores that there are other conceptions of law and indeed that developments illustrated by this case intensify pressures on the legal system to move away from rule-based, hierarchical conceptions. Second, it ignores that these developments are also transforming social work. They are pushing social work away from its traditional proclivity toward informal and tacit judgment toward a style of practice that combines individuation of service with explicitness and standardization of explanation and measurement (Roberts and Yeager 2006).

Key features of the Alabama and Utah decrees can be found in most child welfare reform plans. Such plans virtually always involve commitments to “system investments” or infrastructure development, such as equipment acquisition or training for sophisticated information processing systems; increase of caseworkers and supervisors, with a goal toward getting caseloads down to some benchmark (in foster care, for example, fifteen to twenty cases per worker); and establishment of minimum qualifications for workers and supervisors (for example a bachelor’s degree in social work for frontline workers).

In addition, the agency typically commits to monitor compliance with a series of procedural and documentation norms. Compliance is reviewed

---

7. Paul Vincent described the Alabama decree and Leecia Welch, lead counsel on the Utah case, described the Utah decree in such terms in conversations with us.

through a desk audit of a sample of cases. The Utah process involves about fifty “case processing” benchmarks. For example: Did the worker see the child (reported as being abused) within the priority time frame (from an hour to three days depending on the urgency of the report)? If the child remained at home, did the worker initiate services within thirty days of the referral? If the case involved an allegation of medical neglect, did the worker obtain an assessment from a health care provider prior to case closure? (Utah Division of Child and Family Services 2004).<sup>8</sup> We noted such norms in command-and-control decrees like *G. L.* They still play a role in Alabama and Utah but a less central one.

Three other features of the Alabama and Utah approaches, though hardly unique, are developed with distinctive emphasis: (1) a repudiation of rule-bound authority in favor of a contextual understanding of norms; (2) a distinctive understanding of the relation between the administrative center and local units; and (3) an incrementalist approach to reform.

A statement of core principles is a standard feature of most decrees, but these statements often give way to *G. L.*-type regimes of rigid specification. In contrast, in Alabama and especially Utah, the authoritative pronouncements emphasize that practice should be driven directly and continuously by principles. “Instead of specifying the precise means of accomplishing these ends,” the Alabama decree begins, “[this] decree lays out a set of ‘operating principles’ or ‘standards’ and directs defendant to ensure that the Alabama Department of Human Resources child protective systems comply with them” (*R. C. v. Walley* 1991, Consent Decree, 2). The system emphasizes to workers “the larger context of their decisions” (Utah Division of Child and Family Services 1999, 9). Utah styles its basic norms as “guidelines.” Each section of the guidelines is preceded by a statement of the underlying “philosophy” of the norms. A few norms, mostly concerning process and documentation, take a categorical form, but even they permit “documented exceptions,” where the worker can explain and substantiate the reason for deviation.

The underlying idea—that norms must be both learned and elaborated in the course of practice—is reflected in the reconfigured training that turned out to be central in both states. A large part of traditional casework training involved review of abstract psychological theory about such matters as child-parent “attachment” without any clear delineation of the practical implications of the theory. The more clinical dimension tended to focus on rules and compliance, typically taking the workers through discrete sets of routines for separate stages of a case, from intake to closure. Training under the reforms focuses on a basic skill set that is portrayed as central to every phase of casework: “engagement” (interviewing and counseling), “teaming” (deliberation), “assessment” (diagnosis), “planning,” and “tracking” (follow-up). The

---

8. Data tracking on reporting on some of these measures is federally required (Government Accountability Office 2003).

skills are taught in the context of actual or hypothetical cases through written exercises or role-playing. At the core of each skill is a capacity for contextual judgment.

The second general theme of these experimentalist decrees is a distinctive conception of the relation of central administration and the frontline. In this view, the center articulates general goals, provides support for the frontline, and monitors its success in achieving the goals. Frontline offices and workers have relatively broad discretion to apply governing principles to particular cases. As the Utah plan asserts, “if front-line staff members are focused on correct priorities, and if the administration provides necessary resources and removes barriers to getting things done, the child welfare system can meet its mission of protecting children” (8). This idea entails, in addition to a new style of monitoring we discuss below, fiscal decentralization that gives local offices control over resources and the capacity to contract for goods and services they deem necessary (rather than, as before the reform, limiting them to a menu of standard options provided by centralized procurement) (Bazon Center for Mental Health Law 1998, 26–60). The central administration also provides technical assistance to local offices to help them build expertise in areas where the offices are deficient.

Incrementalism is a third general theme. The initial plans in *G. L.* and New Jersey contemplated an immediate transformation of the entire system. Alabama and Utah adopted phased implementation, decreasing the risk that the center would be overwhelmed by the demands of reform. For example, Alabama set a goal of converting about a seventh of its counties to the new model each year until all had converted. Counties competed to participate in the first cohort based on evidence of strong leadership and effective collaboration with schools, community groups, and social service agencies. Successful competitors received additional funds and staff positions, waivers from rules impeding experimentation, and extensive technical assistance from various consultants. In the course of training in these counties, workers were encouraged to reimagine practice in a way that would vindicate the goals of the reform. The pilot counties created local models that inspired and instructed the ones that followed (30–31). Utah initiated reform across the state at the same time, but in each of the five regions, it evaluated progress separately, and it permitted separate exit from the most intensive monitoring when a region met certain milestones.

### **Reconfiguring Practice**

At a more concrete level, the contribution of the Alabama-Utah approach lies in the way it gives practical structure to three longstanding themes of the background legislation—customization of service, collaborative decision making, and most important, diagnostic monitoring. In the latter



area, the reforms have produced the QSR—which has important general implications for the rule-of-law issues with which we are concerned.

Although it is useful to speak of the Alabama and Utah reforms in terms of a single model, the model does not have a canonical definition. Many participants see the core of the reform as the reconception of frontline case work as contextual and collaborative judgment, sometimes called in Utah and Alabama “the Practice Model” (and elsewhere, the “problem-solving model”) (Huntington 2006, 674–87). Others put greater weight on central facilitation of diagnostic monitoring, especially through the QSR. Despite these differences, there is a consensus that both elements are crucial. What we call the Alabama-Utah model is a heuristic that explains how the integration of collaborative casework with diagnostic monitoring makes it possible for administration to learn from local practice while correcting its mistakes.<sup>9</sup>

### *Customization*

Long-term casework starts with “assessment”—an effort to diagnose the key obstacles to achieving the core child welfare goals. As the Utah plan puts it, “the Practice Model will emphasize the search for underlying causes of the incidents that threaten the safety, permanence, and well being of children” (Utah Division of Child and Family Services 1999, 9). A training vignette illustrates assessment with a case involving the “Archuleta family” that begins with a report to the agency that young children have been left alone at night (US Department of Health and Human Services). The parents eventually return home under the apparent influence of drugs, so the children are placed with a maternal great aunt. The father, whose drug use constitutes a parole violation, is reincarcerated. After extensive investigation, analysis identifies the following needs: (1) the children need a permanency plan (returning home seems possible; long-term care with the great aunt seems a possible alternative); (2) both parents have drug use problems; (3) the mother needs to increase her employment prospects, so she can replace the income lost from the father’s reincarceration; (4) the mother feels socially isolated; (5) the parents need to maintain contact with each other after the father’s incarceration; (6) the parents need greater understanding of child

---

9. We do not mean to suggest that this ambition is unique to Alabama and Utah. Vincent and Groves have contributed to efforts in about a dozen states to implement variations of their model, and other states have adopted analogous reforms independently (see Service Improvement/Systems Reform at the Child Welfare Information Gateway, <http://www.childwelfare.gov/systemwide/service> (accessed July 1, 2008)). Alabama and Utah seem distinctive because of the strong emphasis on diagnostic monitoring in the form of the QSR, because of the comprehensiveness and articulateness of their vision, and because of the relatively disciplined and thorough implementation of the vision.

development and parenting skills; (7) one of the children has delayed speech; (8) both children are anxious about the disruption of their family life; and (9) the great aunt/caregiver may need some time for herself (in addition to the time the children are in school) and needs transportation to get the children to and from school and health care appointments.

Training and review encourages workers to look beyond surface indications, using available diagnostic tools and data in a systematic manner. For example, reviewers in a recent Utah case review faulted the assessment of the child's difficulty in school for failure to consider reports of educational testing in the school records or to follow up on signals in the health records that the child had suffered oxygen deprivation at birth (Utah Division of Child and Family Service 2006a, 12). In another case that was discussed as exemplary at a recent Utah review (one from a randomly selected sample), the caseworker obtained a grant from a private charity to pay the expenses of the nurse on the case team and the foster mother to travel out-of-state to a conference on fetal alcohol syndrome to help determine whether to pursue this issue with regard to two children. The diagnosis was eventually confirmed and treatment initiated. (The foster mother later adopted the children.)<sup>10</sup>

Assessment leads to "planning." Intervention has to be tailored to the circumstances of each child. This means, to begin with, the basic principle, established by AACWA, that removal to foster care should not be the only or even the presumptive response to family dysfunction. Services should be provided when they obviate removal and should be tailored to local circumstances. "Services must be adapted to class members and their families; class members and their families must not be required to adapt to inflexible, pre-existing services" (*R. C. v. Walley* 1991, Consent Decree, 18).

The service plan for the Archuleta family includes drug testing and treatment for both parents, a visitation plan providing daily phone contact and supervised visitation between children and parents, grants to the great aunt/caregiver for daycare and transportation for the children, an arrangement with the children's godparents to provide babysitting relief for the great aunt/caregiver, an agreement by the godparents to provide moral support for the mother's drug treatment efforts, weekly counseling at school for the children, speech therapy for one of the children, and English-as-a-second-language education for the mother.

Workers and providers are expected to adjust services according to client needs. For example, if traditional parenting classes are not sufficiently focused on a parent's needs, workers should consider retaining a consultant who can coach the client individually, perhaps in her home (Bazelon Center for

---

10. Each of us participated as a reviewer on Utah QSR teams during 2006 and 2007. Otherwise unattributed statements about casework in Utah come from observations during those experiences.

Mental Health Law 1998, 54).<sup>11</sup> Customization can also involve logistical convenience. In a recent Utah case, when the caseworker and the father could not find a parenting class compatible with the father's work schedule, they obtained the course materials, and the caseworker tutored the father.

A good plan must attempt to anticipate contingencies. From the beginning, the caseworker should have a "long term view" that envisions the desired conclusion of the case and recognizes the "transitions" that are likely to occur on the way there. A plan for a teenager who is expected to remain in foster care should include preparation for when she "ages out" of the system.<sup>12</sup> If a parent is to be released from jail in a year, the plan should address the team's position on how custody should be affected and what visitation there should be.

At the same time, past judgments have to be reconsidered in the light of new experience, and adjustments have to be made to unanticipated contingencies. For example, in a case involving a teenager who had been put on medication for attention deficit disorder (ADD) with apparent success, the team decided to cut back the medication to see if it was still necessary. When ADD symptoms returned, they restored the original dosage. This was considered good "tracking" practice. On the other hand, the worker in the same case was faulted for failing to learn that the child's school performance had fallen off in the past couple of months, following a family crisis precipitated by an injury to the foster father, and to reassess in the light of this contingency.

By insisting on flexible and tailored responses, the "practice model" strives to attenuate the contrast between the "preservation" and "rescue" approaches to child welfare. It does so by developing a range of possibilities between leaving the child in the home and terminating parental contact. In cases where there is doubt about the possibility of family reunification, workers engage in "concurrent planning," providing support with a view to reuniting the family, but at the same time developing fallback options in the

---

11. For another example, consider the description of part of the service plan for a morbidly obese girl: "A local council agreed to sponsor Jeannie and the homemaker [whom the agency hired to help Jeannie and her mother with nutrition issues] for Weight Watchers. The Christian Service Center provided diet food, treats, and drinks. A beautician agreed to provide periodic styles and perms as rewards for losing weight. Other businesses donated clothing, toys, and a bike, which Jeannie began riding when [the homemaker] wasn't available to take walks with her. Jeannie's mother began coming home from work early to prepare meals, instead of relying on the grandmother" (Bazelon Center for Mental Health Law 1998, 66).

12. "The long term view in this case is optimal. It is clearly understood by all of the team members and contains specific written actions, goals, and milestones. [The youth] has plans to complete high school and attend college at the University of Utah. She sees herself living in the dorm but was able to discuss the possibility that remaining with [the current caregiver] may be a more reasonable start. Funding sources are a shared knowledge of the team, and the worker and [the youth] have a plan whereby she can receive help in purchasing a car. Part of her long-term view is learning how to manage her time, think through her decisions, manage her money, use a checking account and cook and clean. All of the team members are helping her develop these skills" (Utah Division of Child and Family Services 2006b, 26).

event that proves impossible. Even where there is no prospect of reunification, continued parental contact often proves possible and desirable (39).

### *Collaboration*

A central part of the caseworker's job is to "engage and facilitate a child and family team to support the child/youth and family including the out-of-home caregiver and community resources" (Utah Division of Child and Family Services n.d., section 300.4(c)). The child, if over ten years old (Alabama) or twelve years old (Utah), is regarded as a "partner" in his or her plan, as are the natural parents unless they refuse or there are extreme circumstances. Foster parents, as well as friends and relatives whose participation is desired by the family and foster parents, are welcome.

Practice norms speak of parents as both right-holders and sources of information and support. Involving parents and giving them a sense of "ownership" in the plan is an important caseworker duty. Review scores for "system performance" are reduced when parental participation is unsatisfactory for whatever reason. If necessary to permit parents to attend, team meetings should be scheduled outside of work hours or at a parent's home.

The other dimension of collaboration involves professionals. In the traditional image of professionalism—exemplified in law by the judge and in social work by the caseworker—the process of professional judgment is individual and partly tacit; the professional is expected to draw on background experience in a way that cannot be fully articulated. In the reform practice model, decisions are collaborative and explicit. Key judgments are made by a team, and the team is so cognitively diverse that its members must often articulate assumptions that would remain unstated in more homogeneous settings.

The team typically includes the caseworker, who should have a social work or related degree, a health professional who monitors medical services, two lawyers (the child's guardian *ad litem* and an assistant attorney general who represents the state), and a therapist. A teacher, tutor, or school tracker might also join.

An example of the kind of reframing such collaboration is designed to encourage was observed in a recent debriefing session among QSR reviewers. The case involved an autistic teenage boy who was large and strong and was likely to be perceived as threatening or disruptive when agitated. The reviewers responsible for the case explained their high evaluation of "team coordination" by noting that the team had worked with the school to develop a way of physically restraining the boy when he became upset and that school personnel had learned how to perform these "take downs" without any apparent long-term adverse effects. The narrative offered in support of this judgment indicated that the boy had been restrained in school on five recent

occasions. However, another participant in the review session who had had experience with older autistic children questioned whether such frequent restraint was necessary. He noted that disruptive episodes occurred less frequently at home than in the school, even though the child spent more time at home, and his caretaker was not, formally at least, as versed in the treatment of autistic children as the school personnel. He suggested that either the school was too quick to restrain or that it was insufficiently attentive to the conditions that caused the boy to become agitated. Since the case team had not considered or investigated these possibilities, their work was reevaluated less positively.

### **Monitoring**

Utah and Alabama use a variety of monitoring approaches, including the type of “case processing” and aggregate outcome data emphasized in the early *G. L.* and New Jersey decrees and required in some instances by federal regulations.

However, the reformers use a distinctive monitoring procedure—the QSR—designed to complement the customizing and collaborative features of their practice model. The QSR was first applied in child welfare in Alabama and has since been applied to child welfare programs in eleven other states, including Utah. In both the Alabama and Utah lawsuits, the QSR became the central measure of compliance in decisions to terminate court supervision.

The “fundamental assumption” of the QSR is that “each case is a unique and valid test of the system” (Child Welfare Policy and Practice Group n.d. (b), 3). The QSR preserves the traditional social work commitment to forms of supervision that respect the complex contextuality of frontline decisions and encourage workers to respond to clients as concrete individuals. The reformers contrast the “audit focus” of conventional monitoring to the “qualitative” QSR approach:

AUDIT FOCUS: “Is there a current service plan in the file?”

QUALITATIVE FOCUS: “Is the service plan relevant to the needs and goals and coherent in the selection and assembly of strategies, supports, services, and timelines offered?” (1).

Yet, the QSR also aspires to generate precise judgments that can be compared across cases and scores that can be aggregated.

The basic QSR process begins with a stratified random sampling of cases. In Utah, the annual QSR review samples seventy-two cases for Salt Lake City (out of a total population of about fifteen hundred), and twenty-four cases for

the four less populated regions. In Alabama's county-run system, samples range from sixty-eight cases for the largest county to twelve for the smaller ones. Samples are adjusted so that each office has at least one review, and no worker has more than one, and so that there is a balance of in-home and out-of-home interventions, older and younger children, and boys and girls.

The cases are reviewed by teams of two. Ideally (but not always in practice) new reviewers take a training curriculum that involves an explanation of scoring measures and practice case vignettes. They get feedback on their scores until their judgments align with those of veterans. At least one team member must be experienced; the second can be an initiate. In monitoring regimes under judicial decrees, often one reviewer is the monitor or an agent of the monitor and the other is an agency official.

Case reviews take about two days. They start with a file review and then proceed to interviews with the child, family members, nonfamily caregivers, professional team members, and others (for example, teachers) who might have relevant information. The interviews are informal but structured by the basic norms of assessment and individualized planning that guide primary casework.

However, the reviewers must ultimately score the case numerically in terms of indicators. The current instrument has twenty-one. Eleven concern "child and family status": they measure the well-being of the client and his or her family over the past thirty days. The other indicators concern "system performance" over the past ninety days. The two sets of indicators respond to the goals of what business scholars Robert Kaplan and David Norton (1996) call "the balanced scorecard":

A good Balanced Scorecard should have a mix of outcome [status] measures and [system] performance drivers. Outcome measures without performance drivers do not communicate how the outcomes are to be achieved. They also do not provide an early indication about whether the strategy is being implemented successfully. Conversely, performance drivers—such as cycle time or part-per-million defect rates—without outcome measures may enable the business to achieve short-term operational improvements but fail to reveal whether the operational improvements have been translated into . . . enhanced financial performance. (150)

The principal QSR instrument is a ninety-five page booklet with a series of suggestive questions for each indicator (Human Systems and Outcomes 2005). (For example, for "Stability": "How many out-of-home placements has this child had in the past two years? For what reasons? Of the placement changes, how many have been planned? How many have been made to unite the child with siblings/relatives, move to a less restrictive level of care, or make progress toward the permanency goal?" (18).)

Scoring for individual indicators is on a six-point scale with 6 optimal and 1 the worst; 4 is considered minimally acceptable. For aggregation purposes, some indicators receive more weight than others. Weighted average scores of 4 or higher for “child status” and “system performance” indicates that the case as a whole is minimally acceptable. However, the “safety” indicator is a trump for aggregation purposes; a case can be scored as acceptable only if safety is acceptable. After preliminary scoring, the reviewers meet to discuss their cases and particularly to surface and resolve any issues of scoring.

Each case produces a summary like Table 1.

Reviewers meet with the caseworker and supervisor to discuss their findings and the scores. Caseworkers can appeal the scoring through a series of procedures that leads to central management, but this rarely happens.

**TABLE 1.**  
**Child Status and System Performance Ratings**

| Child Status                              | Rating | System Performance                       | Rating |
|---|--------|--|--------|
| 1a. Safety of the Child                   | 5      | 1. Child and Family Participation        | 5      |
| 1b. Safety Risk to Others                 | 5      | 2. Child and Family Team/Coordination    | 4      |
| 1. Overall Safety                         | 5      | 3. Child and Family Assessment           | 4      |
| 2. Stability                              | 5      | 4. Long-Term View                        | 5      |
| 3. Appropriateness of Placement           | 6      | 5. Child and Family Planning Process     | 5      |
| 4. Prospect of Permanence                 | 4      | 6. Plan Implementation                   | 6      |
| 5. Health/Physical Well-Being             | 6      | 7. Formal & Informal Supports & Services | 5      |
| 6. Emotional/Behavioral Well-Being        | 4      | 8. Successful Transitions                | 5      |
| 7. Learning Progress (5 and older)        | 3      | 9. Effective Results                     | 4      |
| 8. Developing/Learning Progress (under 5) | n/a    | 10. Tracking and Adaptation              | 5      |
| 9. Caregiver Functioning                  | 5      | 11. Caregiver Support                    | 5      |
| 10. Family Functioning & Resourcefulness  | n/a    | 12. OVERALL PERFORMANCE                  | 5      |
| 11a. Child Satisfaction                   | 5      |  |        |
| 11b. Parent/Guardian Satisfaction         | 5      |  |        |
| 11c. Substitute Caregiver Satisfaction    | 6      |  |        |
| 11. Overall Satisfaction                  | 5      |  |        |
| 12. OVERALL STATUS                        | 5      |  |        |

When cases are aggregated, a summary for the system or subsystem looks like the one in Table 2 from the 2006 QSR for the Utah Northern Region (Utah Division of Child and Family Services 2006a, 4–6).

The reviewers discuss among themselves the indicators and information from their cases that might explain their significance, and they meet collectively with personnel from the region to discuss the diagnostic importance of their findings. The final report sets out the aggregate scoring, generalizes about what appear to be recurring problems, and presents illustrative examples from specific cases.

The QSR serves three general functions. First, it is a form of clinical training for the caseworkers and their supervisors. The experience of presenting an actual case in its particularity and receiving critical feedback on it from peers and mentors is the core form of professional development in the social work tradition (Munson 1979). Accounts in Alabama and Utah suggest that the QSR has had major effects in changing the frontline experience of supervision and review in ways that have improved both learning and morale. A lawyer team member in one of the Utah cases we reviewed said, “The case workers used to think of quality assurance as a way for the central office to dump on them. Now they think of it as a way for them to show how good they are” (interview, January 24, 2007).

Second, the QSR process is a form of norm elaboration through peer review that engages all levels of the system, as well as outside experts. The meaning of adequacy with regard to goals like safety and permanence or practices like assessment and planning is, in the abstract, indeterminate. The QSR is a collaborative process for specifying norms through analysis of cases. The scoring system forces the reviewers to formulate their judgments in ways sufficiently precise to permit comparisons across cases. The various discussions among reviewers aimed at “inter-rater reliability” and the exchanges between reviewers and frontline workers promote convergent understandings of how the standards apply in particular cases.

For example, an issue that arose early in Utah was whether violence by the father against the mother should be presumed a threat to the safety of the child (and whether a recurrence of spousal abuse should be deemed a recurrence of child abuse). Consensus soon emerged that, if it occurs in the presence or with the knowledge of the children, it should be so presumed.

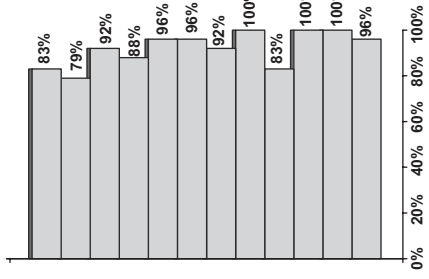
State central administration officials who participate in reviews across the state promote consistency across regions. The integration of outsiders from other states or consultants with national practices promotes consistency across states. Even where consistency is not achieved—there are currently different views among states as to when spousal abuse is tantamount to child abuse—discussion serves to surface issues for further consideration.

Third, QSR data functions as a measure of performance and as diagnostic tool of systemic reform. The scores can be compared over time and (in principle though not as yet in practice) across states. Moreover, they give



TABLE 2.

| Overall System           |                | Northern System Performance |   |      |      |      |      |      | FY07 Current Scores |
|--------------------------|----------------|-----------------------------|---|------|------|------|------|------|---------------------|
|                          | # of cases (+) | # of cases (-)              | Exit Criteria 70% on Shaded indicators/Exit Criteria 80% on overall score |      |      |      |      |      |                     |
|                          |                |                             | FY03  | FY04 | FY05 | FY06 | FY07 | FY08 | FY09                |
| C&F Team/Coordination    | 20             | 4                           | 42%   | 67%  | 75%  | 71%  | 83%  | 83%  |                     |
| Functional Assessment    | 19             | 5                           | 42%   | 54%  | 60%  | 54%  | 79%  | 79%  |                     |
| Long-term View           | 22             | 2                           | 25%   | 58%  | 71%  | 75%  | 92%  | 92%  |                     |
| C&F Planning Process     | 21             | 3                           | 46%   | 63%  | 79%  | 83%  | 88%  | 88%  |                     |
| Plan Implementation      | 23             | 1                           | 71%   | 71%  | 83%  | 88%  | 96%  | 96%  |                     |
| Tracking & Adaptation    | 23             | 1                           | 67%   | 71%  | 88%  | 83%  | 96%  | 96%  |                     |
| C&F Participation        | 22             | 2                           | 50%   | 88%  | 96%  | 67%  | 92%  | 92%  |                     |
| Formal/Informal Supports | 24             | 0                           | 75%   | 79%  | 96%  | 92%  | 100% | 100% |                     |
| Successful Transitions   | 19             | 4                           | 63%   | 73%  | 83%  | 82%  | 83%  | 83%  |                     |
| Effective Results        | 24             | 0                           | 75%   | 71%  | 96%  | 92%  | 100% | 100% |                     |
| Caregiver Support        | 13             | 0                           | 94%   | 92%  | 92%  | 92%  | 100% | 100% |                     |
| Overall Score            | 23             | 1                           | 58%   | 79%  | 83%  | 88%  | 96%  | 96%  |                     |



| Overall Status             |                |                |                                    |             |             |            |            |                     |  |
|----------------------------|----------------|----------------|------------------------------------|-------------|-------------|------------|------------|---------------------|--|
| Northern Child Status      |                |                |                                    |             |             |            |            |                     |  |
|                            | # of cases (+) | # of cases (-) | Exit Criteria 85% on overall score | FY03        | FY04        | FY05       | FY06       | FY07 Current Scores |  |
| Safety                     | 24             | 0              | 100%                               | 100%        | 100%        | 96%        | 96%        | 100%                |  |
| Stability                  | 20             | 4              | 83%                                | 79%         | 75%         | 92%        | 75%        | 83%                 |  |
| Approp. of Placement       | 24             | 0              | 100%                               | 100%        | 96%         | 96%        | 100%       | 100%                |  |
| Prospects for Permanence   | 21             | 3              | 88%                                | 42%         | 67%         | 71%        | 71%        | 88%                 |  |
| Health/Physical Well-being | 24             | 0              | 100%                               | 100%        | 100%        | 100%       | 100%       | 100%                |  |
| Em./Beh. Well-being        | 22             | 2              | 92%                                | 88%         | 79%         | 75%        | 92%        | 92%                 |  |
| Learning Progress          | 22             | 2              | 92%                                | 79%         | 75%         | 83%        | 92%        | 92%                 |  |
| Caregiver Functioning      | 14             | 0              | 100%                               | 88%         | 100%        | 100%       | 100%       | 100%                |  |
| Family Resourcefulness     | 14             | 3              | 82%                                | 44%         | 56%         | 76%        | 71%        | 82%                 |  |
| Satisfaction               | 22             | 2              | 92%                                | 75%         | 92%         | 100%       | 96%        | 92%                 |  |
| <b>Overall Score</b>       | <b>24</b>      | <b>0</b>       | <b>100%</b>                        | <b>100%</b> | <b>100%</b> | <b>96%</b> | <b>96%</b> | <b>100%</b>         |  |

C&F = Child and Family  
 Em./Beh. = Emotional and Behavioral

rough but serviceable indications of where attention and remedial effort should be focused.

Everyone recognizes that aggregate QSR data is crude as a measure of system performance. Because of the cost of reviews of this intensity, it is rarely plausible to get a large enough sample size for statistical validity, though the statistical deficiencies of QSR data are mitigated by the availability of other data. QSR data can be checked against and combined with quantitatively denser data. Moreover, informal information gained in the review process can add systemic perspective. Once the QSR process surfaces a problem and makes it a focus of discussion, frontline caseworkers or managers may have a sense of how pervasive the problem is.

More fundamentally, for purely remedial or diagnostic purposes, precise systemic generalization is not necessary. For these purposes, it is more important to spot problems and remedy them than to characterize the current state of the system with statistical precision. Here the QSR designers share premises with modern business management perspectives like “Total Quality Management” with its “zero tolerance” for major defects (Crosby 1984, 74–84). The system should strive for superior performance in every case, and every substantial failing should be regarded as a problem worth remedying. Thus, the designers insist that their sampling goal is to get a sufficient number of cases that any systemic problem will show up in at least one of the cases reviewed. They have no formula for ascertaining this number, but it seems likely to be considerably lower than the one required for statistical validity at high confidence intervals. To be sure, if the QSR process discloses a large number of problems, some prioritization may be necessary, and it may then be important to have more precise measures of systemic pervasiveness. But after the system has crossed a threshold of proficiency, it may not be necessary to worry about the relative systemic magnitudes of the problems disclosed.

A particular diagnostic focus in analysis of aggregate QSR results has been comparison of in-home or protective service cases to foster care cases. Indications that children in one group are faring worse than another are treated as signals to reassess whether adequate and proper services are being provided. For example, QSR scores for both status and system were significantly lower for in-home than foster children in Utah’s Western region for several years. The region focused training efforts on perceived problems in these cases, and the difference disappeared in the 2006 review (Utah Division of Child and Family Services 2006c, 35–36).<sup>13</sup>

---

13. In some cases, it appears that workers were underestimating the danger of recurrence of abuse from natural parents. The administrators decided that the “permanency” workers assigned to the cases had insufficient experience with the kind of serious abuse problems most often encountered at the initial investigation stage by “protective service” workers. The response was to assign protective service workers to consult routinely with permanency workers on in-home cases (interview with Linda Winiger, Director of Practice Improvement, Utah DCFS, March 19, 2007).

## Systemic Assessment in Public Law Litigation: Termination in Alabama and Utah

In January 2007 the district court granted the defendant's motion to terminate the decree in Alabama, and in May 2007 the parties agreed to a phased termination of the decree in Utah. Both actions reflect a conception of welfare rights that sees process, rather than rule-compliance or outcomes, as primary.

The Alabama decision shows how QSR data can be combined with more narrowly quantitative data and more impressionistic information to assess performance (*R. C. v. Walley* 2007). The decision squarely rejects rule-based conceptions of compliance. The district court refused to treat as dispositive the state's extensive failure to meet the requirement of the 1997 federal legislation that termination of parental rights be initiated after fifteen months in care, asserting that such a "rigid time guideline" is not feasible or desirable (48). Instead, the court looked more generally at the state's efforts to achieve "permanency," as measured by the QSR, and it declined to consider even this more general value in isolation from others. (An isolated focus on permanency encourages a rush to get children into placements that are long-term but not necessarily optimal.)

The most important measure of compliance was the benchmark, agreed to by the parties, of 85 percent of cases with an overall score of no less than "acceptable" under a QSR review supervised by the monitor. Each of the state's sixty-seven counties had met the benchmark at least once in recent years, but some had subsequently regressed. In the monitor's most recent review of ten representative counties, three had not met the benchmark.

Turning to more specific indicators of the QSR and case processing data, the court found benchmark compliance across a substantial range but also significant failures. With respect to each failure, the court asked two questions. First, is there some mitigating explanation for the failure? For example, the court accepted the defendant's explanation that a large part of the increase in out-of-home placements reflected an epidemic of methamphetamine use that data showed accounted for as much as 70 percent of intakes in some counties.

---

Another example: Utah analysts found statewide that status scores in a major subcategory of foster care cases—where the children are placed with relatives—were lagging other subcategories. Administrators responded by focusing additional review efforts on kinship placement cases. They concluded that decision makers had been giving too strong a preference to relatives in placement decisions. They also found training and support for kinship placements had been inadequate. For training and support purposes, workers had been treating relative caregivers more like natural parents than like unrelated foster parents. Natural parents, on average, require less training and support because they have experience with their own kids. Training efforts were reconfigured to sensitize workers to the problems of kinship placements and practice guidelines were revised to increase the presumptive training and support to relative foster parents, with noticeably improved results (interview with Richard Anderson, August 18, 2006).

Second, the court asked, did the agency independently identify the failure and initiate plausible corrective action? With respect to methamphetamine cases, the agency had inaugurated collaboration with law enforcement agencies. With respect to excessive caseloads, the state had developed procedures to track caseloads closely and to reassign cases when necessary. It had responded to high worker turnover by increasing pay scales, and it responded to deficiencies in planning with increased training and an effort to simplify the planning process.

The court's analysis concluded with a brief assessment of "sustainability." It approved the agency's plan to retain much of the monitoring regime constructed under the consent decree, and it noted that the regime issues "report cards" published on its Web site for each county. It also mentioned that the department will continue to be subject to review by the federal Department of Health and Human Services and a governor's "child welfare commission."

The court's decision to terminate the decree is certainly debatable. The opinion glosses over serious deficiencies in performance in the largest county—Jefferson—where the city of Birmingham is located and which has 20 percent of the state's caseload.<sup>14</sup> Its discussion of how quality assurance and other accountability mechanisms are likely to operate after termination is brief and superficial; other than the reference to the report cards, there is no discussion of transparency or mechanisms for continuing collective participation by stakeholder groups.

Nevertheless, the court's framing of the critical inquiry is plausible. It interprets the welfare rights claims in the case to entail not the judicial elaboration and enforcement of substantive standards, but the inducement of the parties to construct a process that will elaborate and implement them. The key features of this process are transparency and accountability to the public and stakeholders and a capacity for self-assessment and self-correction.

In contrast to the situation in Alabama, the parties in Utah agreed that termination was warranted. Like the Alabama district court's decision, the termination agreement in Utah focuses on the system's internal capacity for diagnosis and adjustment, and it addresses issues of external accountability more extensively than the Alabama decision. In the agreement, the state

---

14. The monitor concluded that "the data do not support a finding that the county is sustaining performance in accordance with the R. C. principles and the five core purposes" (*R. C. v. Walley* Consent Decree, n.d., 88). On the QSR, overall system performance was adequate in 57 percent of cases. Child status was adequate in 97 percent of cases, but "child/family satisfaction" was an exceptionally low 53 percent. About one-third of workers had excessive caseloads. Nearly one-third of frontline positions had turned over in the preceding year, and length of training had recently been reduced. In addition, the use of congregate (institutional) care had increased "dramatically," the percentage of siblings placed together had decreased, and the percentage of children in care for thirteen months or longer had increased from 46 percent to 65 percent (79–89).

promises to “take steps to protect the integrity of the [QSR] process, as necessary” (*David C. v. Huntsman* 2007, paragraph 19). It further commits to maintain until at least the end of 2010 the key elements of what we have called the Alabama-Utah model. These include the “practice model” (the training and supervision associated with customization of services and collaborative judgment), the QSR, review of quantitative case processing data, and a general capacity for “self-critique and analysis to improve practice” (paragraphs 18–30, Appendix B).

With respect to external accountability, the agreement contains several provisions requiring that performance data, including QSR data, and reports be published online. It requires various periodic reports to the public, the legislature, and nongovernmental organizations (NGOs). It also requires increased support for statewide and regional Quality Improvement (QI) committees. The QI committee members are appointed from public agencies, NGO partner organizations, local businesses, and professional groups with relevant expertise. They receive staff support and are charged with review of trends and recommendations of reform. One provision of the agreement specifically guarantees them the opportunity to participate as QSR reviewers (paragraphs 31–35).

### III. COMPLETING THE SYSTEM: FEDERAL REVIEW AND EMPIRICAL RIGOR

The Alabama-Utah model is not a complete system of accountability. It operates at the state level: a complete system requires national oversight. Moreover, assessment of the efficacy of practice in the Alabama-Utah model is short term and informal. Ideally, practice should be evaluated in terms of more long-term and systematic data. However, the Alabama-Utah approach can serve as a model for its federal complement. And while it does not itself amount to scientific validation, it provides an important constituent of any effort to achieve it.

#### Federal Oversight

In the experimentalist view of American democracy, federal administration plays an oversight and facilitation role by holding the states accountable, measuring their performance, and pressuring them to improve. At the same time, it assists them with technical and material resources and facilitates the exchange of information among them (Dorf and Sabel 1998).

Federal oversight of state child welfare efforts has yet to vindicate the aspirations of this view. However, in recent years, the Administration of Children and Families (ACF) of the Department of Health and Human

Services has developed an assessment process—the Child and Family Services Review (CFSR)—that shows great promise. Despite limitations in its current implementation, the CFSR has the potential to become an important national complement to the innovative state reforms. The CFSR, launched in 2000 and revised several times since, has evolved toward increasing resemblance to the Alabama-Utah model (in part, under the influence of Alabama) (Center for the Study of Social Policy 2003, 28–29). It combines review of aggregate case processing and outcome data with qualitative peer review of a sample of cases. It has become less punitive in orientation and more diagnostic and remedial. Key federal officials see “continuous improvement” as the core CFSR value (Christian 2001; National Child Welfare Resource Center for Organizational Improvement 2007).

In states that did not already have qualitative monitoring processes, the CFSR review has prompted their initiation. States that already had such review have tailored it to complement the federal process. In some cases, they have used QSR-type data to challenge the findings of federal CFSR review, initiating the kind of mutual interrogation that is part of the promise of federalism.

The CFSR remains a flawed process. For example, its audit sample is small even by QSR standards, its performance measures blur process and outcome in a way that impairs their diagnostic value, and its results are not compiled or disclosed in a way that facilitates comparison across states (National Coalition for Child Protection Reform 2003). Most importantly, the CFSR currently selects its own cases for review, while a more plausible procedure would re-review a sample of the cases from the state’s process. The federal process should be auditing the state’s audit rather than conducting an independent one; that way, its efforts leverage the states’ and integrate with them directly. Yet, these flaws seem remediable, and the trend of federal practice is promising.

### **Filling Empirical Lacunae**

Even when amplified by case processing data and federally required reporting, the QSR has two important limitations as a mode of social service assessment. First, it does not control for the severity or difficulty of cases. This impairs the value of comparisons of performance scores for different systems or the same systems over time. Different scores could reflect differences in the average conditions and needs of children in different systems, or in the harshness of the social environments in which the systems operate, or they could reflect differences in the system’s intake practices. A system that assumes custody of children at a relatively low threshold of danger will find it easier, on average, to reunify or achieve stability for children than a system that intervenes only in clear cases of danger.

Second, QSR data says little about the long-term effects of practices. It provides a snapshot of current practice and the current status of children in care. Thus, it usefully identifies both failure to comply with practice norms and practice norms that are ineffective in the short term. But it does not say much about the long-term robustness of practice or how well it develops enduring capacities in its charges.

A fully elaborated system would deal with both limitations. It should generate baseline estimates of the average degree of challenge involved in a caseload that can be used to adjust performance scores for longitudinal and cross-section comparisons. It should also attempt to systematically assess the long-term efficacy of its practices.

It is important to note that such efforts, far from obviating QSR-type processes, would depend on them heavily. Estimating a baseline of caseload difficulty requires that cases be rated on relevant parameters in terms of a uniform standard. It would be easy to integrate some such rating into the process by which initial intervention decisions are made. However, for such ratings to be plausible and consistent, they would have to be subject to a review process with many of the characteristics of the QSR. Extending the QSR to initial intervention decisions would be an important component of an effort to generate a baseline of case difficulty.

QSR-type methods can also contribute to long-term assessment. The paradigmatic method for such assessment, familiar from clinical drug trials, is the random assignment controlled experiment. A key difficulty of such studies in the social service field is the specification and control of the interventions being studied.<sup>15</sup> You cannot measure the efficacy of a program unless you can see the extent to which it is actually being implemented. The QSR is the best means of doing so.

#### IV. THE ANTINOMIES OF WELFARE RIGHTS

Discussion of the rule of law in the welfare system has clustered around four antinomies: (1) rules versus standards; (2) bureaucratic versus adjudicatory control; (3) narrow versus broad judicial intervention; and (4) negative versus positive rights. The Alabama-Utah model suggests responses to each that moderate the tensions between their opposed terms.

---

15. For a discussion of methodology and an illustrative study, see Wald, Carlsmith, and Liederman (1988). Abhijit Banerjee (2007) recently criticized economic development studies that measure returns to investment in, for example, education without looking at what teachers are actually doing or even whether they show up for class. Education is one of many areas in which QSR-type processes could play an important role in assessment and improvement of complex services.



### (1) Rules versus Standards

The Progressive and New Deal visions of welfare administration accorded broad discretion to frontline workers operating under general standards. Both conservative and liberal critics in the 1960s and 1970s found pervasive abuse of discretion and urged that workers be hemmed in by highly specified rules (Platt 1977; Simon 1984).

A salient example was the formalization of decision making about employment availability in the Social Security Disability program. An eligibility condition for the program is that the applicant be incapable of performing any job that exists in significant numbers in the national economy. Prior to 1978, this issue was decided through an informal all-things-considered judgment. The Social Security Administration, responding in part to the cost of this practice and in part to concerns about the inconsistency across different decision makers, produced a “grid” that mechanically dictated answers on the basis of a limited number of factors, such as applicant’s level of education and physical strength (*Heckler v. Campbell* 1983).

Yet, recent decades have seen an increasingly prominent countertrend toward more individuation (Minow 1990). For example, statutes governing benefits to developmentally disabled people, children in need of “special education” services, and indigent families now require individual plans with customized services. While the Social Security Disability grid remains in place, Jerry Mashaw (2006) reports that dissatisfaction has led some to urge that the system move toward “a community-based and multidisciplinary approach that would deploy financial assistance, medical care, rehabilitation, and transportation services, among other things, to promote the overall well-being and highest possible functioning of disability beneficiaries.” This new approach, Mashaw emphasizes, “would demand highly discretionary judgments” (154–55).

Mashaw portrays this return to standards, discretion, and individuation as a transient episode in a cyclical oscillation between categorical and contextual norms (see also Lipsky and Smith 1989). Another view, however, sees the trend as more fundamental and secular. Surveying developments in Europe, the Irish National Economic and Social Development Office (2005) sees individuation, or what it calls “tailored universalism,” as a key theme of an emerging “developmental welfare state” (203). Its analysis emphasizes that recent social and economic change has upset traditional premises of European and American welfare systems. Increased geographical mobility and immigration has made the populations served by welfare programs more diverse. Core beneficiaries of traditional welfare programs—women and the elderly—have been increasingly pushed and pulled into the labor market, requiring that the programs intended for them be redesigned to better accommodate the mixing of public support and employment. Economic development has increased the vulnerability of the less skilled segments of the workforce,

calling for transitional public support that combines income transfers and training.

However, this latter account has not explained how individuating, adaptive programs can respond plausibly to the problem of administrative discretion. The Alabama-Utah model aspires to provide an answer. Child welfare jurisprudence in Alabama and Utah is resolutely standards-oriented in the sense that interpretation is purposive; specific norms are understood as expressions of general higher-level norms, and interpreters of all ranks feel directly responsible to the highest-level purposes of the system.

Yet decision makers feel pervasively accountable. Accountability arises, in part, from the duty to explain decisions to others in various review procedures, notably the QSR. This duty mandates, in the first place, articulation. Decision makers do not expect deference to ineffable wisdom attested to by professional credentials. Good case work is not considered adequate if it is not consistent with and derived from a written plan, with accompanying documentation of its premises. A common complaint within the practice model is that “the plan is not driving the practice.”

The emphasis on articulation serves several purposes. It facilitates outside review of the team’s work, and it makes it easier for new members of the team to acquire understanding of the team’s prior work. Equally important, it facilitates the team’s internal functioning. Having to articulate their views forces each member to think them through as clearly as possible. Having to agree on a common formulation increases the chance that they share a common understanding.

Rules do have a role in this system, but it is not the role contemplated in the traditional jurisprudence that gives rise to the antinomy of rules and standards. Their role there is to restrict discretion. In the model we are considering, discretion is restricted by qualitative peer review and by public reporting of monitoring results. The function of rules is to make explicit the learning that results from the review process. Once review makes clear some gap between the core purposes of the system and their expression in some lower-level rule, the rule gets revised so that its guidance is clearer.

The simultaneous emphasis on articulation and flexibility appears to increase the congruence of norm and conduct in comparison with command-and-control regimes. The old systems were governed formally by relatively restrictive rules, but in practice the rules were easily ignored. Paul Vincent says of the Alabama situation, “Frankly, the practice-related content of the manual got little attention . . . prior to the decree. Caseworker practice had been governed more by worker bias and the local office practice culture than State office influence” (Child Welfare Policy and Practice Group n.d. (a), 5).

The resolution of the rules/standard antinomy in the Alabama-Utah model has a striking manifestation—it seems to have muted the century-old debate between the family preservation and rescue models. The academic literature on child welfare remains dominated by these competing

perspectives. Yet, the debate is not salient among the participants in the Alabama and Utah reforms.

The most important practical stake in the debate concerns what presumption or default rule will be applied in situations where our information about a case is incomplete or inconclusive. In situations of doubt, should we remove or preserve? The more ignorant or uncertain we are, the more important the presumption. Thus, Congress and policy makers, remote from the circumstances of particular cases and distrustful of frontline workers, have devoted great energy to the presumption mandated by AACWA.

But it seems that when cases are looked at with the contextual intensity and professional care that AACWA prescribes and Alabama and Utah frequently achieve, the default norm is rarely important. With surprising frequency, professionals and stakeholders can agree on what the “best interests” of the child are without invoking any background theoretical or ideological premise.

For example, in the first few years of the new regime in Alabama, decision makers came to recognize that some of their family reunification efforts were not productive. In these cases, they were providing elaborate services over long periods, but the family situation was not improving. As Paul Vincent remembers, “workers had gotten into the mindset of thinking that every removal was a failure.”<sup>16</sup> The QSR system identified these cases, and administrators responded with corrective discussion, training, and guidelines. This reorientation occurred long before Congress’s 1997 reformulation of the AACWA reunification presumption to reduce what the legislature then perceived as excessive commitment to family preservation. In a system with the capacity for careful contextual response and adaptation, the presumption was relatively unimportant.

## (2) Bureaucratic versus Adjudicatory Control

The historical oscillation between rules and standards is paralleled by a tension between managerial and adjudicatory review of frontline decisions.

On the one hand, the rule of law assumes that rules will be misapplied and that misapplications will be corrected by adjudication. The Social Security Act of 1935, for example, required that beneficiaries be afforded opportunities for hearings in both its social insurance and public assistance (AFDC) programs, and the Supreme Court later held that such hearings are often a matter of constitutional right (*Goldberg v. Kelly* 1970). The hearing right entails an opportunity of the beneficiary to present evidence and arguments to an officer with some measure of independence from the original decision maker.

---

16. Interview with Paul Vincent, August 24, 2006.

Yet, it eventually became apparent that the hearing process was not adequate to accomplish the broader rule-of-law goals of cases like *Goldberg*. For one thing, beneficiaries lacked the knowledge and resources to identify legally questionable decisions and appeal them. Although rates of hearing decisions for claimants were often high, only a tiny fraction of frontline decisions are taken to hearing in most programs. There have been many indications that large numbers of decisions that could have been reversed at hearings went unchallenged. Moreover, in the few areas where appeal rates are substantial, such as Social Security Disability and immigration asylum, there appears to be broad inconsistency among hearing officers; comparable cases are likely to be decided differently depending on which officer decides them (Simon 1984; Ramji-Nogales, Schoenholtz, and Schrag 2008).

Such considerations led lawyers to argue that there was a “management side of due process” (Mashaw 1974). Thus, the welfare lawyers argued for “quality assurance” systems that would protect beneficiary interests by auditing representative samples of frontline decisions, remedying the specific errors uncovered, and taking systemic remedial action when the audits disclosed patterns of error. Moreover, however adequate hearing procedures might be to protect beneficiary interests, they did not even purport to protect nonbeneficiary public interests in accurate decision making and efficient administration. Thus, conservatives joined the call for more bureaucratic control.

Such concerns produced an array of audit and quality control processes in the 1970s and 1980s. These efforts, however, were in turn disappointing, especially to those concerned with beneficiary interests. The quality assurance systems typically paid more attention and applied severer sanctions to errors that favored beneficiaries than to errors that harmed them. Moreover, they tended toward the kind of rule-bound “checklist” review that rigidified practice in ways many perceived as counterproductive of the goals of the programs (Simon 1984; Lipsky 1984). The literature on “street-level bureaucracy” suggested that the activities of frontline public workers were effectively unreviewable except perhaps for conformity to rigid rules that only tenuously expressed the important public values at stake (Lipsky 1980; Wilson 1968).

In Alabama and Utah, the QSR motto, “Every case is a unique and valid test of the system,” expresses the model’s ambition to transcend the dichotomy between bureaucratic and judicial control. The QSR combines the features of intensive concern with individual interests and stakeholder participation characteristic of adjudication with the systemic review associated with bureaucracy. The idea is that, in a system committed to radical individuation, intensive review of the particularities of an individual case is the most important mode of systemic diagnosis. Cases are learning opportunities, and review provides diagnostic feedback.

There are two key systemic dimensions of the QSR process. One is the process by which individual case data is aggregated and generalized to achieve indications of systemic problems. We have seen that the QSR instrument

itself is a kind of “balanced scorecard” that links outcome and practice indicators in ways that invoke causal hypotheses about system efficacy. In addition, individual QSRs can provoke informal “root cause analysis” in which decision makers reflect on the systemic inferences of problems disclosed in the cases. And QSR scores are analyzed with more quantitative case processing data.

The other systemic dimension is a process by which reviewers, decision makers, and administrators deliberate to achieve consistency in scoring cases. The process aims at what social scientists call “inter-rater reliability,” and as in social science, it is achieved through deliberation.<sup>17</sup> Unlike judges in the traditional view, QSR reviewers do not do their jobs “one case at a time,” and they do not make their isolation from the systems they review a point of pride. Unlike bureaucrats in the traditional view, they do not achieve consistency through rules. Rather, they try to generalize across cases by sharing perspectives and experiences in ways that strive for consistency without sacrificing particularity.

There are, of course, important roles for adjudication, as more or less conventionally conceived, in child welfare. Internally, child welfare systems need administrative adjudicatory processes to respond to stakeholder complaints. To date, such processes seem to play a minor role in most jurisdictions, most likely because they have not been ambitiously implemented. In addition, such processes are not well articulated with the processes of systemic review like the QSR. Ideally, such systems should be related to QSR-type processes, so that the information they generate can be combined with QSR data and other diagnostic indicators.

Externally, traditionally independent courts are needed, first, to decide intractably contested high-stakes issues, such as custody and termination of parental rights, and second, to intervene structurally in response to persistent, systemically inadequate administrative performance. To varying degrees, the courts seem to perform these roles relatively well, and they could perform the

---

17. The review teams seem well structured to avoid the twin dangers of deliberative processes—stalemate, on the one hand, and “groupthink,” on the other (Sunstein 2000). In principle, the members share general familiarity with and commitment to the basic goals of the program, but they have sufficient cognitive diversity that tendencies to unreflective agreement are restrained.

It is possible that either through unreflective drift or bad faith calculation, the review teams might turn out to be monolithically and uncritically sympathetic toward the agency in a manner analogous to the portrayal by Edelman, Uggens, and Erlanger (1999) of the evolution of workplace grievance mechanisms. However, there are several potential checks on this tendency, including: (1) the federal CFSR, which will audit or reaudit state cases with reviewers of its own choosing; (2) the “citizen’s panels” required by federal statute, which could plausibly perform their role by participating in and monitoring the state’s auditing practices (including their choice of reviewers); and (3) stakeholder groups without direct authority, but who will often be able to credibly threaten reform litigation if the agency’s efforts at self-assessment are demonstrably insincere or ineffective.

first better if they were relieved of much of the burden of routine supervision they must now perform when administrative systems malfunction severely.

### (3) Discrete versus Systemic Judicial Intervention

Lon L. Fuller (1978) raised doubts about the role of courts in the welfare system by suggesting that “polycentric” claims were ill suited for judicial intervention. Polycentric problems arise in complexly integrated systems where a judicial mandate with respect to one part might ramify in unpredictable or uncontrollable ways to other parts.

Justice Black expressed a variation on this concern in his dissent in *Goldberg v. Kelly* (1970, 278–80). If courts require welfare programs to afford pretermination hearings, he speculated, the programs are likely to respond by making it more difficult to establish eligibility in the first place. In fact, the course of administrative reform in the 1980s validated this worry. Documentation and verification requirements were increased, so that the process became more burdensome, and a large fraction of applications came to be rejected on procedural grounds (Lipsky 1980). And one can even find the converse phenomenon. When a court ordered New York’s special education program to improve its processing of eligibility determinations, it shifted staff away from providing services to existing beneficiaries, and service to them declined (Sandler and Schoenbrod 2003).

Polycentricity calls for systemic intervention, but systemic intervention involves a different problem. The courts despair of deriving and enforcing determinate norms for the conduct of an entire system. Such concerns sometimes lead them to hold even statutory mandates of systemic practice judicially unenforceable (*Pennhurst State School v. Halderman* 1981; *Suter v. Artist M.* 1992).

However, the kind of judicial intervention exemplified by Alabama and Utah is a response to polycentric problems that does not require the court to analytically derive a system on its own. The Alabama-Utah approach enables the court to induce an agency that has persistently failed to meet its responsibilities to reform in collaboration with injured stakeholders in a way that is both accountable and transparent. Far from imposing the kind of rigid and arbitrary regime that some critics of structural remedies fear, it inhibits administrative rigidity and arbitrariness by inducing the agency to develop its capacities to assess itself and respond to experience.

Paradoxically, there is a sense in which the broadest remedies are the least intrusive. The ideal structural decree requires no more than what good management and democratic accountability would require in the absence of judicial intervention. The Alabama-Utah approach comes much closer to this ideal than its command-and-control predecessors. In cases that have taken this direction, there are fewer complaints from defendants that

court-imposed monitoring is a burdensome distraction. Indeed, the Utah plan said that it was “first and foremost, a business plan which the Division intends to implement voluntarily with or without further court involvement” (Utah Division of Child and Family Services 1999, 5).

One indication that broad intervention can be less intrusive is that defendants sometimes collaborate with plaintiffs in order to broaden classes to avoid the arbitrary effects of narrow intervention. The holdings of *Suter v. Artist M.* (1992) and cases that preceded it, denying a private right to enforce AACWA, were intended to limit the possibilities of broad intervention. With AACWA unavailable, the most salient alternative grounds were the substantive due process claim of *Youngberg v. Romeo* (1982), which was predicated on state custody and the antidiscrimination norm of the Rehabilitation Act, which was limited to people with disabilities. Remedies limited to children in custody or with diagnosed disabilities would have created pressure to take custody of children simply in order to make them eligible for needed services or to give them diagnoses of mental health disabilities for the same reason. Such pressures had been experienced in the North Carolina *Willie M.* case, where the class was limited to children in custody with mental health problems (Dodge, Kupersmidt, and Fontaine n.d.). Mindful of this experience, plaintiffs and the defendant in Alabama, first, agreed to include in the class both children in custody with mental health problems and children “at risk” of being taken into custody and developing mental health problems and, second, stipulated that *all* children at risk of being taken into custody were at risk of developing mental health problems (Bazelon Center for Mental Health Law 1998, 16–17). Confining the court to narrow intervention constrains its discretion only at the cost of arbitrariness.

#### (4) Negative versus Positive Rights

The canonical statement of the priority of negative rights in American constitutional law—*DeShaney v. Winnebago Department of Social Services* (1989)—arose in the child abuse-and-neglect area. The Supreme Court held that the reckless failure of the state to protect a child from severe parental abuse did not constitute actionable deprivation of “life, liberty, and property” under the civil rights laws.

The traditional objection to positive constitutional rights is that the relevant principles are indeterminate. Theorists contend that welfare programs are not underpinned by a body of evolving but specifiable social norms comparable to those that give coherence to judicial decision making about private rights. Welfare systems lack the self-adjusting properties of private markets; they have to be steered by bureaucracies under political supervision. Thus, judicial intervention along traditional rule-of-law lines disrupts

political accountability and threatens rigidity or arbitrariness or both (Hayek 1944; Dennis and Stewart 2004; Teubner 1986).

To the extent that the indeterminacy claim is true, it implies a terrible trade-off. Either we must exempt from the strongest rule-of-law protection some of the most basic and important interests of a broad fraction of the population, or we must empower or burden the courts with the task of defining and enforcing standards that are not susceptible to coherent judicial elaboration.

Child welfare provides an especially striking example of the arbitrariness of the distinction between negative and positive rights in any effort to impose rule-of-law values on the welfare state. Perhaps the least controversial proposition in the field is that a state's failure to provide services to troubled families increases the likelihood that children will be removed. It strikes many as incoherent to give priority to a negative right against unnecessary child removal over a positive right to services that would obviate removal.

Fortunately, the more promising child welfare cases and recent public law litigation in other areas suggest a conception of public right that eludes the distinction between negative and positive rights and a mode of enforcement that could be applied as readily in situations of wrongful state action and situations of wrongful failure to act (Sabel and Simon 2004).

At the most general level, the core welfare right is an entitlement to have one's interests in a public program considered in a process that is responsive and accountable. The "interests" in question are those relevant to the core values of the program. "Responsive" implies consideration of the relation of the claimant's interests to the relevant public purposes and at least minimal participation by the claimant. "Accountability" implies a reasoned explanation by the decision makers, review of decisions in ways that provide rich assessments of both individual cases and the system as a whole, and transparent procedures of systemic self-assessment and self-correction<sup>18</sup> (Handler 1986).

The *prima facie* case of a violation of this type of public law right is a showing of, first, the state's chronic failure to meet relevant standards of performance in the area and, second, immunity of the system to conventional forces of political correction. The remedy that follows from a finding of liability is not a judicially imposed code, but a judicial order that the system negotiate and implement with the claimants and other stakeholders a reform program that is accountable in the ways suggested by the Alabama and Utah reforms.

---

18. The accountability connoted by the rule of law runs in several directions: (1) accountability to individual claimants; (2) accountability of lower-level decision makers to upper-level decision makers; (3) accountability of the agency as a whole to the other branches of government, stakeholder communities, and the general public. The QSR contributes directly and powerfully to the second and third kind of accountability. It also contributes to the first, but it needs to be supplemented by processes in which claimants can initiate review.



The shadow of such a conception of social rights can be found in the holding of the Supreme Court in *Youngberg v. Romeo* (1982) that due process requires certain welfare decisions to be based on “professional judgment” (322). *Youngberg* is impaired by its limitation to rights that can be styled as “negative” (for example, on the facts of the case, rights premised on state custody of the plaintiff) and its undeveloped conception of “professional judgment.” But it seems to contemplate something like the prima facie case we propose—failure to meet basic standards and political closure to key stakeholders. And the remedial practice in the cases brought under it, we have seen, is gradually converging along the lines of Utah and Alabama.

Other relevant doctrinal touchstones are the celebrated decisions of the South Africa Constitutional Court in *South Africa v. Grootboom* (2000) and *Minister of Health v. Treatment Action Campaign* (TAC) (2002). *Grootboom* held that the state could not constitutionally facilitate the eviction of a large group of homeless people when it had failed to make reasonable provision for people in their situation in its subsidized housing programs. *TAC* held that the state acted unconstitutionally in withholding antiretroviral drugs from pregnant women and nursing mothers with HIV. The cases are notoriously ambiguous, but a key theme in *Grootboom* is political exclusion, and a key theme in *TAC* is a violation of basic standards. These are the elements of our public law right. In both cases the remedy—an order to the state to reform its program in ways that make “reasonable” provision for the interests of the claimants—seems consistent with American public law litigation. The developments we have reported suggest how such mandates might be elaborated.

This is not the place for a full-scale account of welfare rights, but we should take note of some ways in which the Alabama-Utah experience responds to the conventional objections to the justiciability of positive rights.

The first objection is that such rights cannot be coherently specified. The objection is confirmed by our cases to the extent that it suggests that courts cannot specify detailed rules for running welfare programs, but it is refuted to the extent that it suggests that such specification is needed for effective intervention.

What courts need to be able to do to enforce public law rights is, first, to determine when programs are chronically and fundamentally failing to meet their commitments and, second, to trigger and facilitate experimentalist reform. Structural reform cases in child welfare and elsewhere strongly support the claim that they are capable of these tasks. Liability determinations in these cases are rarely even contested after the motion phase, and few critics of these cases deny that the systems are broken. Note that legal determinacy in these cases does not depend on constitutional or statutory rules. The specific doctrinal basis of the claims seems to have little influence on the trajectory of intervention. Substantive consistency in liability determinations arises from the standards of “professional judgment” that the courts, in the manner of *Youngberg*, apply to assess agency performance. And

it seems to be arising in the formulation of remedies from mutual learning across cases about the efficacy of alternative structures.

The second objection is that the enforcement of positive rights requires the reallocation of scarce resources at the expense of usually unidentifiable but potentially equally important interests. We have not undertaken the conceptually and practically difficult task of assessing whether and to what extent the Alabama and Utah reforms increased expenses. They clearly required increased appropriations in the short term. In the long term, however, many would expect the reforms to reduce expenditures by decreasing the number of removals to foster care (which tends to be more costly than natural family support), minimizing the use of congregate care (in general the most expensive intervention), and speeding the time to “permanency.”<sup>19</sup>

Moreover, states typically find that good management increases access to outside funding. A key component of most reform plans is the reorganization of efforts to maximize federal reimbursement under Medicaid, Supplemental Security Income, and the child welfare programs of the Social Security Act. Creative local efforts often draw money from churches and foundations. To the extent that competitively awarded grant money is drawn to reformed programs—because they seem to offer a higher social return—the objection that positive rights enforcement cannot take account of the relative values of competing uses seems wrong.

The most fundamental point, however, is that experimentalist intervention, if successful, is not zero-sum and need not, in principle, cost anything. In business, a prominent slogan in an influential school of management literature is, “Quality is free” (Crosby 1984). No doubt it is an exaggeration, but it expresses the basic insight that value can be created through reorganization.

The third criticism is that the courts lack either the political legitimacy or the political power to mandate broad scale public reforms. With respect to legitimacy, consider that experimentalist reform can induce judicial accountability in the same way that it induces administrative accountability. The mechanisms of review, assessment, and transparency that courts impose on defendants can be used to assess the legitimacy and efficacy of the courts’ efforts. Appellate courts can assess the plausibility of the parties’ statement of relevant legal standards, of the appropriateness of the indicators they have chosen to measure their performance, and what improvement these measures show.

Experimentalist reforms also facilitate legislative oversight by making program performance more transparent. The Utah experience also shows a

---

19. A key difficulty in assessing the increase in expenses due to the intervention would be constructing the baseline of what expenses would be in the absence of intervention. To do so would require controlling for variations in the social forces that affect family dysfunction.

further distinctive innovation in the possibilities of legislative oversight. More than a dozen state legislators have participated in the QSR process as reviewers.<sup>20</sup>

As for political power, the Alabama experience shows the importance of political support, but it also suggests that political support is not exogenous to the remedial regime. The Alabama reforms were initially supported by the agency leadership, which assisted the plaintiffs in developing support from the governor and the legislature. Subsequently, a new governor was elected on an antiwelfare backlash platform that explicitly proposed to undo the reforms and oppose continued federal court intervention. The new regime did serious damage to the reforms but not nearly as much as it sought to do. The dismantling efforts were vigorously opposed by a coalition of NGOs, by prominent newspaper editorialists, and by some legislators. Among the key actors in the campaign to save the reforms was a statewide Parents Support Network. The network was formed under the reforms as a forum for moral support and information exchange, but it also played an important role in lobbying the legislature for adequate resources (Bazelon Center for Mental Health Law 1998). The effectiveness of the opposition to the new governor suggests that successful experimentalist reform may generate its own political support. It also suggests that reforms may organize otherwise vulnerable stakeholder constituencies in ways that make the reforms politically more stable.

## CONCLUSION

Alabama and Utah have produced a distinctive model of child welfare administration that incorporates some longstanding reform prescriptions and adds an innovative form of diagnostic monitoring. We have reported impressionistic evidence that the model is responsible for dramatic improvement in both states and is becoming influential elsewhere. More systematically, we have argued that the model gives new perspective to longstanding debates about legal accountability in the welfare state and illustrates a form of accountability of frontline decision making that does not entail the rigidification of practice. It seems readily generalizable to other service-based welfare programs and especially responsive to the increasing tendency of such programs to aspire to adaptability and customization. Most generally, it suggests ways in which the conception of “positive” or “social” constitutional rights hinted at in some prominent recent discussions might be elaborated.

---

20. Interview with Richard Anderson, August 18, 2006. Legislative oversight is often thought to be limited to “police patrol” procedures that audit for compliance with fixed rules and “fire alarm” procedures that respond to reports of discrete abuses (McCubbins and Schwartz 1984). The QSR suggests an alternative that could provide far richer information.

## REFERENCES

- ABA Center on Children and the Law. 2005. *Child Welfare Consent Decrees: Analysis of Thirty-Five Court Actions from 1995 to 2005*. Chicago: American Bar Association.
- Banerjee, Abhijit. 2007. Inside the Machine: Toward a New Development Economics. March/April, *Boston Review*.
- Bartholet, Elizabeth. 1999. *Nobody's Children: Abused Children, Foster Drift, and the Adoption Alternative*. Boston: Beacon Press.
- Bazon Center for Mental Health Law. 1998. *Making Child Welfare Work: How the R. C. Lawsuit Forged New Partnerships to Protect Children and Sustain Families*. Washington, DC: Bazon Center on Mental Health Law.
- Center for the Study of Social Policy. 2003. Improving the Performance and Outcomes of Child Welfare Through State Improvement Plans (PIPs). [http://www.cssp.org/uploadFiles/2515\\_CSSP\\_FINAL.pdf](http://www.cssp.org/uploadFiles/2515_CSSP_FINAL.pdf) (accessed July 1, 2008).
- Chayes, Abram. 1976. The Role of the Judge in Public Law Litigation. *Harvard Law Review* 89:1281–316.
- Child Welfare Policy and Practice Group. n.d. (a). Implementation of Alabama's R. C. Consent Decree: Creating a New Culture of Practice. [http://www.childwelfaregroup.org/documents/AL\\_RC\\_Implementation.pdf](http://www.childwelfaregroup.org/documents/AL_RC_Implementation.pdf) (accessed February 22, 2009).
- Child Welfare Policy and Practice Group. n.d. (b). The Qualitative Service Review Process. <http://www.hspolicy.utah.gov/dcfs/> (accessed July 1, 2008).
- Christian, Steve. 2001. Child and Family Services Reviews. NCSL State Legislative Report. National Conference of State Legislatures. <http://www.ncsl.org/programs/cyfl/slr265.htm> (accessed February 1, 2009).
- Costin, Lela B., Howard Jacob Karger, and David Stoesz. 1996. *The Politics of Child Abuse in America*. New York: Oxford University Press.
- Crosby, Philip B. 1984. *Quality Without Tears*. New York: McGraw-Hill.
- Dennis, Michael J., and David P. Stewart. 2004. Justiciability of Economic, Social, and Cultural Rights. *American Journal of International Law* 98:462–515.
- Dodge, Kenneth, Janis Kupersmidt, and Reid Fontaine. n.d. *Willie M.: A Legacy of Legal, Social, and Policy Change*. Durham, NC: Duke University (Center for Child and Family Policy).
- Dorf, Michael, and Charles Sabel. 1998. A Constitution of Democratic Experimentalism. *Columbia Law Review* 98:267–473.
- Edelman, Lauren, Christopher Uggen, and Howard Erlanger. 1999. The Endogeneity of Legal Regulation: Grievance Procedures as Rational Myth. *American Journal of Sociology* 105:506–54.
- Fuller, Lon L. 1978. The Forms and Limits of Adjudication. *Harvard Law Review* 92:358–409.
- Garrison, Marsha. 1987. Child Welfare Decisionmaking: In Search of the Least Drastic Alternative. *Georgetown Law Journal* 75:1745–828.
- . 2005. Reforming Child Protection: A Public Health Perspective. *Virginia Journal of Social Policy and Law* 12:590–637.
- Guggenheim, Martin. 2005. *What's Wrong With Children's Rights?* Cambridge, MA: Harvard University Press.
- Government Accountability Office. 2003. *Most States Are Developing Statewide Information Systems, but Reliability of Child Welfare Data Could Be Improved* (GAO 03-809, July). Washington, DC: Government Accountability Office.
- Handler, Joel. 1986. *The Conditions of Discretion: Autonomy, Community, Bureaucracy*. New York: Russell Sage Foundation.

- . 2004. *Social Citizenship and Workfare in the United States and Europe: The Paradox of Inclusion*. Cambridge: Cambridge University Press.
- Hanson, Chris. 1994. Making It Work: Implementation of Court Orders Requiring Restructuring of State Executive Branch Agencies. In *Child Parent, and State: Law and Policy Reader*, ed. S. Randall Humm, 224–36. Philadelphia: Temple University Press.
- Harlow, Carol. 1999. Accountability, New Public Management, and the Problems of the Child Support Agency. *Journal of Law and Society* 26:150–74.
- Hayek, Friedrich. 1944. *The Road to Serfdom*. Chicago: University of Chicago Press.
- Human Systems and Outcomes. 2005. Quality Service Review: Protocol for Use by Certified Reviewers. In authors' possession.
- Huntington, Clare. 2006. Rights Myopia in Child Welfare. *UCLA Law Review* 53:637–700.
- Kaplan, Robert S., and David P. Norton. 1996. *The Balanced Scorecard: Translating Strategy Into Action*. Cambridge, MA: Harvard Business School Press.
- Lipsky, Michael. 1980. *Street-Level Bureaucracy: Dilemmas of the Individual in Public Service*. New York: Russell Sage Foundation.
- . 1984. Bureaucratic Disentitlement in Social Welfare Programs. *Social Service Review* 58:3–27.
- Lipsky, Michael, and Steven Rathgeb Smith. 1989. Nonprofit Organization, Government, and the Welfare State. *Political Science Quarterly* 104:625–48.
- Lowry, Marcia. 1999–2000. Foster Care and Adoption Reform Litigation: Implementing the Adoption and Safe Families Act of 1997. *St. John's Journal of Legal Commentary* 14:447–56.
- Mashaw, Jerry L. 1974. The Management Side of Due Process: Some Theoretical and Litigation Notes on the Assurance of Accuracy, Fairness, and Timeliness in the Adjudication of Social Welfare Claims. *Cornell Law Review* 59:772–824.
- . 2006. Accountability and Institutional Design: Some Thoughts on the Grammar of Governance. In *Public Accountability: Designs Dilemmas, Experiences*, ed. Michael Dowdle, 115–56. Cambridge: Cambridge University Press.
- McCubbins, Matthew, and Thomas Schwartz. 1984. Congressional Oversight Overlooked: Police Patrols versus Fire Alarms. *American Journal of Political Science* 28:16–79.
- Minow, Martha. 1990. *Making All the Difference*. Ithaca, NY: Cornell University Press.
- Munson, Carlton, 1979. *Supervision in Social Work: Classic Statements and Theoretical Issues*. New York: Free Press.
- National Child Welfare Resource Center for Organizational Improvement. 2007. Comments on the Second Round . . . An Interview with Jerry Milner, Senior Child Welfare Specialist. <http://muskie.usm.maine.edu/helpkids/rcpdfs/cwmatters5.pdf> (accessed March 11, 2009).
- National Coalition for Child Protection Reform. 2003. The Trouble with Child and Family Services Reviews. <http://www.nccpr.org> (accessed July 1, 2008).
- National Economic and Social Development Office. 2005. *The Developmental Welfare State*. Dublin: National Economic and Social Development Office.
- New Jersey Department of Children and Families. 2004a. A New Beginning: The Future of Child Welfare in New Jersey. July 9. [www.state.nj.us/humanservices/CWA/A%20New%20Begining/A\\_New\\_Begining.pdf](http://www.state.nj.us/humanservices/CWA/A%20New%20Begining/A_New_Begining.pdf) (accessed July 1, 2008).
- . 2004b. Enforceable Elements of the New Jersey Child Welfare Plan. July 19. On file with authors.
- . 2005. New Jersey Period I Monitoring Report, February. [http://cssp.org/uploadFiles/Final\\_NJ\\_Monitoring\\_Report\\_02\\_23\\_07.pdf](http://cssp.org/uploadFiles/Final_NJ_Monitoring_Report_02_23_07.pdf) (accessed March 5, 2009).
- . 2008. New Jersey Period III Monitoring Report, April 16. <http://www.cssp.org/uploadFiles/FINAL%20NJ%20REPORT%20Apr%202008.pdf> (accessed March 5, 2009).

- Platt, Anthony. 1977. *The Child Savers*, 2nd ed. Chicago: University of Chicago Press.
- Ramji-Nogales, Jaya, Andrew Schoenholtz, and Philip Schrag. 2008. Refugee Roulette: Disparities in Asylum Adjudication. *Stanford Law Review* 60:295–411.
- Roberts, Albert R., and Kenneth A. Yeager. 2006. *Foundations of Evidence-Based Social Work Practice*. New York: Oxford University Press.
- Ross, Catherine. 2003. The Tyranny of Time: Vulnerable Children, Bad Mothers, and Statutory Deadlines in Parental Termination Proceedings. *Virginia Journal of Social Policy and Law* 11:176–228.
- Sabel, Charles F. 1995. Learning by Monitoring: The Institutions of Economic Development. In *Handbook of Economic Sociology*, ed. Neil Smelser and Richard Swedberg, 135–65. Princeton, NJ: Princeton University Press.
- Sabel, Charles F., and William H. Simon. 2004. Destabilization Rights: How Public Law Litigation Succeeds. *Harvard Law Review* 117:1015–101.
- Sandler, Ross, and David Schoenbrod. 2003. *Democracy by Decree*. New Haven, CT: Yale University Press.
- Simon, William H. 1984. Legality, Bureaucracy, and Class in the Welfare System. *Yale Law Journal* 92:1198–269.
- . 2006. Toyota Jurisprudence: Legal Theory and Rolling Rule Regimes. In *Law and New Governance in the EU and the US*, ed. Grainne de Burca and Joanne Scott, 37–64. Oxford: Hart Publishing Co.
- Stack, Barbara White. 2005. Law to Increase Adoptions Results in More Orphans. *Seattle Post-Intelligencer*, January 3. [http://seattlepi.nwsourc.com/national/206256\\_orphans03.html](http://seattlepi.nwsourc.com/national/206256_orphans03.html) (page no longer active).
- Sunstein, Cass. 2000. Deliberative Trouble: Why Groups Go to Extremes. *Yale Law Journal* 110:71–120.
- Teubner, Gunther, ed. 1986. After Legal Instrumentalism?: Strategic Models of Post-Regulatory Law. In *Dilemmas of Law in the Welfare State*, 305–13. Berlin: Walter de Gruyter.
- Tushnet, Mark. 2003. Social Rights and the Forms of Judicial Review. *Texas Law Review* 82:1895–920.
- US Department of Health and Human Services. Administration for Children and Families, Comprehensive Family Assessment Process. [www.acf.hhs.gov/programs/cb/pubs/family\\_assessment/fa4.htm](http://www.acf.hhs.gov/programs/cb/pubs/family_assessment/fa4.htm) (accessed March 5, 2009).
- Utah Division of Child and Family Services. n.d. Practice Guidelines and Rules. <http://www.hspolicy.utah.gov/dcfs/> (accessed July 1, 2008).
- . 1999. *The Milestone Performance Plan*. <http://www.hsdcsf.utah.gov/PDF/milestone.pdf> (March 5, 2009).
- . 2004. Case Process Review Questions—2004. Not published. In authors' possession.
- . 2006a. Quality Case Review: Northern Region. Not published. In authors' possession.
- . 2006b. Quality Case Review: Salt Lake City Region. Not published. In authors' possession.
- . 2006c. Quality Case Review: Western Region. Not published. In authors' possession.
- Wald, Michael S., J. M. Carlsmith, and P. H. Liederman. 1988. *Protecting Abused and Neglected Children*. Cambridge, MA: Harvard University Press.
- Wilson, James Q. 1968. *Varieties of Police Behavior*. Cambridge, MA: Harvard University Press.

## CASES CITED

### United States

- Charlie H. V. Whitman*, 83 F. Supp. 2d 476, 483–84 (D.N.J. 2000).  
*David C. v. Huntsman*, Civil No. 2:93-CV-00206TC (D. Utah), Agreement to Terminate Lawsuit, May 11, 2007.  
*David C. v. Leavitt*, 13 F. Supp. 2d 1206 (D. Utah 1998).  
*DeShaney v. Winnebago Department of Social Services*, 489 U.S. 189 (1989).  
*G. L. v. Stangler*, 873 F. Supp. 252 (D. Mo. 1994).  
*G. L. v. Zumwalt*, 564 F. Supp. 1030 (W.D. Mo. 1983).  
*Goldberg v. Kelly*, 397 U.S. 254 (1970).  
*Heckler v. Campbell*, 461 U.S. 458 (1983).  
*Pennhurst State School v. Halderman*, 451 U.S. 1 (1981).  
*R. C. v. Walley*, 2007 WL 118101 (M.D. Ala. 2007).  
*R. C. v. Walley*, Civil No. 88-H-1170-N (M.D. Ala.). Consent Decree. June 5, 1991.  
*R. C. v. Walley* Consent Decree: Monitor's Report to the Court in Response to Court's Order Directing the Monitor to Conduct On-Site Reviews, n.d.  
*Rufo v. Inmates of Suffolk County Jail*, 502 U.S. 367 (1992).  
*Suter v. Artist M.*, 503 U.S. 347 (1992).  
*Youngberg v. Romeo*, 457 U.S. 307 (1982).

### South Africa

- Minister of Health v. Treatment Action Campaign*, 2002 (5) SALR 721 (CC) (S.Afr.).  
*South Africa v. Grootboom*, 2000 (11) BCLR 1169 (CC) (S. Afr.).

## STATUTES CITED

- Adoption and Safe Families Act (ASFA); Pub. L. No. 105–89 (1997), 42 U.S.C. section 1305 et seq.  
Adoption Assistance and Child Welfare Act; Pub. L. No. 96–272 (1980), 42 U.S.C. section 670 et seq.  
Child Abuse Prevention and Treatment Act; 42 U.S.C. section 5101 (1974), 42 U.S.C. section 5101 et seq.  
Rehabilitation Act; Pub. L. No. 93–112 (1973), 29 U.S.C. section 791.  
Social Security Act; Pub.L. No. 74–271 (1935), 42 U.S.C. section 301 et seq.