THE RULE OF LAW IN THE EXPERIMENTALIST WELFARE STATE: LESSONS FROM CHILD WELFARE REFORM

BY:

KATHLEEN G. NOONAN
CENTER FOR HIGH IMPACT PHILANTHROPY, SCHOOL OF SOCIAL POLICY AND PRACTICE, UNIVERSITY OF PENNSYLVANIA

- AND -

PROFESSORS CHARLES F. SABEL & WILLIAM H. SIMON
COLUMBIA LAW SCHOOL

THE RULE OF LAW IN THE EXPERIMENTALIST 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 re-designed 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.

---

[*] Associate Director, Center for High Impact Philanthropy, School of Social Policy and Practice, University of Pennsylvania
[†] Maurice T. Moore Professor of Law, Columbia University
[‡] Arthur Levitt Professor of Law, Columbia University
We are grateful for advice and encouragement to many people, including Richard Anderson, Stephen Cohen, Navina Forsythe, Ivor Groves, Jerry Mashaw, Brad McGarry, George Taylor, Jane Spinak, Paul Vincent, and Michael Wald.

I. Introduction

Recent trends in the evolution of welfare systems in America and abroad intensify longstanding uncertainties about how rule-of-law values apply in these systems. Programs that once focused on financial redistribution increasingly link transfer payments to services. 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. It is driven in part by the perception that these circumstances are more fluid than they have been in the past.

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, first, 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 second, 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.[1] 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.[2]

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 which 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. We find in these developments indications that the institutional innovations that make possible effective customization and adaptation of services also make possible – indeed require – heightened forms of accountability of the systems to their beneficiaries and the larger public. The reforms suggest new theoretical and institutional responses to the dialectics of rule and standard, bureaucratic and judicial accountability, and broad and narrow judicial intervention.

The reforms do not achieve accountability by constraining frontline decisions through rules. Rather, frontline discretion is increased, but joined to the requirement that, in the course of establishing and adjusting plans for children, frontline workers and the professionals and stakeholders with whom they collaborate explain the choices they make in terms of the governing values of the program. Review of these explanations in turn allows administrative superiors and outside oversight bodies to detect and begin considering how to correct misjudgments by individual case workers, systemic flaws in operating routines at the local office or program level, and even ambiguity or mistake in the agency's own conception of its key commitments and plans for achieving them. Thus, the agency learns to improve while monitoring what it does, and the same process that makes

---

[1] See William H. Simon, "Legality, Bureaucracy, and Class in the Welfare System," 92 Yale Law Journal 1198 (1983).

[2] See Abram Chayes, "The Role of the Judge in Public Law Litigation," 89 Harvard Law Review 1281 (1976).

customization of services effective makes it accountable as well. We call such learning-by-monitoring institutions "experimentalist".[3]

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. A series of major federal statutory initiatives failed to impose order on the state-run systems. In at least 30 states, courts have found or defendants have conceded systemic non-compliance with constitutional or statutory requirements on a scale warranting structural intervention.[4] However, some of these interventions have made progress, and the model we find promising has emerged in a handful of them. It may be because these systems have been so deeply broken that they have lent themselves to relatively radical experimentation. And 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 practices with accountability.

Child welfare reform also has 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 system and capable of effective institutionalization across the public sphere. This conception is hinted at in, for example, the welfare rights jurisprudence of the South African

---

[3] Michael Dorf and Charles Sabel, "A Constitution of Democratic Experimentalism," 98 Columbia Law Review 267 (1998). On "learning by monitoring", see Charles F. Sabel. "Learning by Monitoring: The Institutions of Economic Development," in Handbook of Economic Sociology (Neil Smelser and Richard Swedberg, ed.s 1995); William H. Simon, "Toyota Jurisprudence: Legal Theory and Rolling Rule Regimes," in Law and New Governance in the EU and the US (Grainne de Burca and Joanne Scott, ed.s, 2006).

[4] ABA Center on Children and the Law, Child Welfare Consent Decrees: Analysis of Thirty-Five Court Actions from 1995 to 2005 1 (October 2005).

Constitutional Court.[5] It 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 Utah-Alabama model is richly suggestive as to how this notion of rights, only vaguely invoked in the celebrated South African cases, might be elaborated.

Part II elaborates four basic antinomies in modern efforts to apply rule-of-law values to the welfare state. Part III briefly surveys the policy and legal background of the struggle to protect children from abuse and neglect. Part IV elaborates the Alabama-Utah model. Part V discusses processes at the national level that are needed to fully develop the potential of the model. Traditional federal oversight has been erratic and only occasionally effective, but as it absorbs the lessons of Alabama and Utah, it has an important role to play in the scaling the reforms to the national level. Part VI concludes by revisiting the four antinomies of welfare rights in the light of innovations in child protective services.

II. The Antinomies of Welfare Rights

The rule of law connotes basic notions of executive accountability – fidelity to constitutional and legislative authority, consistency in administrative decision-making, and transparency –- from which no one would exempt the welfare system. Moreover, the presumptive mode of enforcement of rule-of-law values in the administrative state – judicial review of administrative action – is well established in modern democracies.

Yet, many continue to doubt that the principles of executive accountability historically developed in connection with private rights can be coherently elaborated in the context of welfare programs. Moreover, there is no consensus among those committed to rule-of-law values in the welfare state as to how those values should be institutionalized there.

These concerns can be seen in four antinomies.

A. Rules v. Standards

---

[5] See, e.g., Mark Tushnet, "Social Rights and the Forms of Judicial Review," 82 Texas Law Review 1895, 1904-09 (2003).

Lawyers are drawn to rules to constrain administrative discretion and promote consistent decision-making. But rule application can be arbitrary in relation to the relevant goals. So lawyers are drawn to standards because they promote individualized consideration of how goals can be vindicated in the context of the particular claimant. The result is an historical oscillation between administrative procedures favoring the one and administrative procedures favoring the other.

The modern American welfare state developed in the early 20th century under the influence of a view that advocated discretion to individualize programmatic responses to the circumstances of the beneficiary. In juvenile courts, education, child protection, and public assistance, the ideal was decision by extensively trained professionals under standards. In the 1960s and 1970s, there was a reaction against this view. Critiques on the right and the left converged in harsh judgments on the performance of the street-level bureaucrats and therapeutic professionals who staffed welfare agencies. They were deemed intrusive, oppressive, and arbitrary.[6]

For many of the critics, the solution involved the control or elimination of discretion through highly specified, strictly enforced rules. Workers were no longer expected or permitted to make complex, professional judgments. They were increasingly confined to check-lists of mechanical routine. In public assistance, for example, the "consolidation" of grants under the Aid to Families with Dependent Children (AFDC) program eliminated the practice of individualized assessment of a household's needs in favor of a "flat grant" varying only with family size. Eligibility workers lost their responsibility to assist clients with a menu of social services and were remitted to policing their compliance with an expanded array of documentation and verification requirements.[7]

A more limited but highly salient development in this trend was the formalization of decision-making about job 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 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 of judgments across different

---

6 E.g., Anthony Platt, The Child Savers (2d ed. 1977)

7 Simon, cited in note , at 1201-06.

decisionmakers, 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.[8]

Yet, recent decades have seen an increasingly prominent counter-trend toward more individuation.[9] For example: Statutes require that educational services to people with learning disabilities be provided in accordance with an "Individualized Education Plan" [10] and that the broader array of services for developmentally disabled beneficiaries be provided so as to "maximize the developmental potential" of each recipient.[11] The Americans With Disabilities Act, in addition to containing a traditional prohibition of discrimination, affirmatively requires employers to make "reasonable accommodation" of disabled employees, a duty that requires individualized assessment of the worker's capacities and needs.[12] The Transitional Aid to Needy Families program that emerged from the Clinton welfare reforms requires an "initial assessment" of employability for all applicants and encourages "personal responsibility plans identifying the education, training, and job placement services needed to move [the recipient] into the workforce." [13]

While the Social Security Disability "grid" remains in place, Jerry Mashaw 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."[14]

On one view, this return to standards, discretion and individuation is a transient episode in an endless oscillation between categorical and

---

[8] See Heckler v. Campbell, 461 U.S. 458 (1983).

[9] See generally Martha Minow, Making All the Difference (1990).

[10] 20 U.S.C. 1400 et seq.

[11] Pub. L. 98-527, sec. 2. The statute also calls for "individualized supports". 42 U.S.C. 15002(8)(A)(v), 15024.

[12] 42. U.S.C. 21112(b)(5)(A).

[13] 42 USC 601 et seq.

[14] Jerry Mashaw, "Accountability and Institutional Design: Some Thoughts on the Grammar of Governance," in Public Accountability: Designs, Dilemmas, Experiences 115, 154-55 (Michael Dowdle ed. 2006).

contextual norms.[15] Another view, however, sees the trend as more fundamental and secular. Surveying developments in Europe, the Irish National Economic and Social Development Office sees individuation, or what it calls "tailored universalism" as a key theme of an emerging "developmental welfare state." 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 re-designed 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.[16]

However, this latter account has not explained how individuating, adaptive programs can respond plausibly to the problem of administrative discretion. In the absence of such an account, a reversion to rules seems most likely.

B. Bureaucratic v. Adjudicatory Internal 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 anticipates that rules can be misapplied and that misapplications 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.[17] 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.

---

[15] See Michael Lipsky and Steven Rathgeb Smith, "Nonprofit Organization, Government, and the Welfare State," 104 Political Science Quarterly 625 (1989); Mashaw, "Accountability," at 155-56.

[16] National Economic and Social Development Office, The Developmental Welfare State 203 (2005); see generally pp. 197-231.

[17] Pub. L. No. 271, sec.s 2(a) (4), 402(a) (4), 49 Stat. 620, 627, 645 (1935); Goldberg v. Kelly, 397 U.S. 254 (1970).

Many programs developed elaborate administrative hearing systems with well-compensated, law-trained quasi-judicial officers housed outside the agency's main line of authority. These officers often developed a professional ethos that valued independence within the agency and respectful treatment of claimants. In many programs, a high fraction of hearing decisions has vindicated beneficiaries.[18]

Yet, it soon became apparent that the hearing process was not adequate to accomplish the broader rule-of-law goals of cases like Goldberg v. Kelly. For one thing, beneficiaries lacked the knowledge and resources to identify legally questionable decisions and challenge them. Although rates of hearing decisions for claimants were often high, the rates of challenge to front-line decisions were small, often tiny. (This was particularly so for public assistance, as opposed to social insurance, programs.) There were many indications that large numbers of decisions that could have been reversed at hearing went unchallenged. [19]

Moreover, a further problem is observed in the one program in which appeal rates for negative eligibility decisions tends to be high (sometimes more than half) – Social Security Disability. There appears to be broad inconsistency among hearing officers; comparable cases are likely to be decided differently depending on which officer decides them.[20]

Such considerations led lawyers to argue that there was a "management side of due process." The most important implication of this idea was adequate legal protection of beneficiaries required efforts to review unappealed decisions by front-line workers.[21] Thus, the welfare lawyers argued for "quality assurance" systems that would protect beneficiary interests by auditing representative samples of cases, remedying the specific errors uncovered, and taking systemic remedial action when the audits disclosed patterns of error.

Such programs were widely implemented in the 1980s but in a way that served fiscal economy goals more than beneficiary protection ones. The quality assurance systems typically paid more attention, and applied severer sanctions, to errors that favored beneficiaries than to errors that

---

18 Simon, cited in note , at 1246-49.

19 Id., and works cited therein at notes 124-26.

20 Id., at 1201-02, 1247-49,

21 See Jerry L. Mashaw, "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," 59 Cornell Law Review 772 (1974).

harmed them.[22] Of course, this bias reflected the politics of program implementation; it is not inherent in managerial control. Nevertheless, there does seem to be an irreducible measure of tension between managerial and adjudicatory approaches. This measure arises from the value of "independence" that is a defining feature of the adjudicatory ideal. Independence implies some immunity from direct supervision and hierarchical accountability. In principle, the adjudicator resolves the claims of the parties before her by applying "the law". To the extent that she appears to be carrying out the instructions of an administrative superior, she no longer seems to be acting judicially. Hearing officers acculturated in this ideal and with a private inclination for autonomy often resist direct administrative supervision.[23]

The main form of accountability that is accepted as consistent with the judicial ideal is party-initiated appeal to a higher judicial body in a process where reversal carries no private adverse consequences for the initial adjudicator. But because they depend on parties with limited knowledge to identify issues and trigger appeals and because of their aloofness from the actors they review, such processes are often inadequate to insure the rule-of-law values of fidelity, consistency, and transparency across the full range of welfare programs.

C. Discrete v. Systemic Judicial Intervention

Lon L. Fuller raised doubts about the role of courts in the welfare system by suggesting that "polycentric" claims were relatively 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.[24]

Justice Black expressed a variation on this concern in his dissent in Goldberg v. Kelly.[25] If courts require welfare programs to afford pre-termination 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 seems to have validated this

---

22 Simon, cited in note , at 1212-13.

23 E.g., Philippe Nonet, Administrative Justice 214 (1969) (reporting that, as the role of California workers' compensation referees was reconceived as judicial rather than administrative, the referees protested supervision by invoking their duty to "decide cases in accordance with our conscience"); Nash v. Califano, 613 F.2d 10 (2d Cir. 1980) (considering challenge by Social Security administrative law judges to supervisory controls as inconsistent with requirements of "decisional independence").

24 Lon L. Fuller, "The Forms and Limits of Adjudication," Harvard Law Review

25 397 U.S. at 278-80.

worry precisely. 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.[26] 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.[27]

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. Thus, in two famous cases, the Supreme Court declined to enforce broad statutory mandates of systemic practice. In Pennhurst State School v. Halderman,[28] the plaintiffs sought injunctive relief to force an institution for the developmentally disabled to comply with the mandate of a federal statute to operate in a manner calculated to "maximize the developmental potential of [each patient] in the setting that is least restrictive of [her] liberty." In Suter v. Artist M[29]., plaintiffs sought injunctive relief to force compliance with the federal Adoption Assistance and Child Welfare Act (AACWA) requirement that a state's child protection program "make reasonable efforts to eliminate the need for removal of the child from his home." The court interpreted each statute as privately unenforceable, in part because of the broad indeterminacy of the standards.

The more narrow and specific the legal mandate, the more the court's enforcement of it threatens to have unforeseeable or undesirable collateral effects. But comprehensive intervention is hard to ground determinately in legal authority.

D. Negative v. Positive Rights

The canonical statement of the priority of negative rights in American constitutional law -- DeShaney v. Winnebago Department of Social Services -- arose in the child abuse-and-neglect area. The plaintiff was a child who suffered severe brain-injuries through repeated beatings by his father. Despite awareness over a long period of extensive evidence of the danger to the child, the state social services agency had intervened only ineffectually and failed to remove him from the home. The complaint

---

[26] Michael Lipsky, "Bureaucratic Disentitlement in Social Welfare Programs," 58 Social Service Review 3 (1984).

[27] Ross Sandler and David Schoenbrod, Democracy by Decree

[28] 451 U.S. 1 (1981).

[29] 503 U.S. 347 (1992).

alleged that this failure constituted a state deprivation of "life, liberty, and property" under the 14th amendment (and hence actionable under section 1983). The court rejected the claim, holding in an opinion by Justice Rehnquist that "the Due Process Clauses generally confer no affirmative right to governmental aid [against lawless private action], even where such aid may be necessary to secure life, liberty, or property interests of which the government itself may not deprive the individual."[30]

For the most part, Rehnquist's opinion gives only history and convention as grounds for the decision. This is the way we do things here, he seems to say. The opinion is thus unsatisfying to those who think that constitutional doctrine should be grounded in principle. The negative/positive distinction does not strongly track any plausible measure of the relative importance of a citizen's interests. This point is illustrated by juxtaposing DeShaney's disqualification of a child's interest in freedom from life-threatening private violence with another Rehnquist opinion finding that a prisoner's property interest in a $23 "hobby kit" inadvertently destroyed by prison employees does qualify for constitutional protection.[31]

However, in its final paragraph, DeShaney does refer briefly to a relevant concern: "In defense of [the defendants] it must also be said that had they moved too soon to take custody of the son away from the father, they would likely have been met with charges of improperly intruding into the parent-child relationship, charges based on the same Due Process Clause that forms the basis for the present charge of failure to provide adequate protection." [32]

Here Rehnquist echoes the long-standing claim that principles of positive right are indeterminate. Theorists contend that government 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.[33]

---

30 DeShaney v. Winnebago Dept. of Social Services, 489 U.S. 189, 196 (1989).

31 Parratt v. Taylor, 451 U.S. 527 (1981).

32 489 U.S. at

33 See, e.g., Friedrich Hayek, The Road to Serfdom (1944); Gunther Teubner, "After Legal Instrumentalism?: Strategic Models of Post-Regulatory Law," in Dilemmas of Law in the Welfare State 299, 305-13 (Gunther Teubner ed. 1986); Michael J. Dennis and David

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.

E. Beyond the Antinomies: Youngberg v. Romeo

Finally, consider a case that seems, in some respects, to stand outside the tensions just rehearsed, and to point toward their possible resolution. Youngberg v. Romeo[34] involved a challenge by child residents of a state institution for the developmentally disabled to conditions at the institution. They alleged that the institution was dangerous, unhealthy, and made no effective efforts to educate its charges. The court held that such conditions could violate the constitutional rights of the residents if they 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."[35] Justice Powell insisted that "it is not for the courts to specify which of several professionally acceptable choices should have been made," but that the constitution did require the courts to "make certain that professional judgment was in fact exercised."[36]

At first glance, the invocation of "professional judgment" may sound like a familiar move within the structure of argument created by the four tensions just elaborated. Professional judgment connotes standards and accountability through adjudication (for example, malpractice actions) rather than bureaucracy. Judicial assurance of the “professional judgment” norm entails broad intervention. At the same time, since the children are in state custody, the rights are negative.

But there is another dimension to the decision that seems to escape the antinomies, a dimension that could be seen as both problem and opportunity. The problem is that no existing profession is an obvious candidate to assume the responsibility for the "professional judgment" on which the court's doctrine hinges. Institutions for the developmentally disabled are run by members of several professions -- medicine, social

---

P. Stewart, "Justiciability of Economic, Social, and Cultural Rights," 98 American Journal of International Law 462 (2004).

[34] 457 U.S. 307 (1982).

[35] Youngberg v. Romeo, 457 U.S. 307, 322 (1982).

[36] 457 U.S. at 319 (quoting Seitz, C.J.).

work, clinical psychology, and education -- whose relative responsibilities have been vague, overlapping, and contested. Moreover, many of the most important judgments made within these institutions do not fit the understanding of professional judgment held within any of these professions. Traditionally, professionals have thought of the judgments their members make as individual decisions made within a single, stable discipline. But increasingly, the key decisions made within social welfare institutions are collective, multidisciplinary, and based on principles that are provisional and changing.

Youngberg spoke as if the "profession" that could vindicate its rule-of-law concerns already existed. In fact, it had to be created. We show below that, in a related area, it is being created. Moreover, it is being created in a way that has the potential to, if not resolve, to relax the four antinomies of welfare rights discourse. The professionalism emerging in the most creative child protection reforms combines rules with standards and centralization and decentralization in novel and promising ways. It makes possible broad judicial intervention in a way that does not require courts to elaborate analytically the operating principles of the programs. And it vindicates rule-of-law values in a way that can be applied as readily on either side of the traditional distinction between negative and positive rights.

III. The Struggle for Child Protection

A. The Federal Response to Child Abuse and Neglect

Abused and neglected children were identified as a social problem at the end of the 19th century by lay philanthropies, most notably a group of local Societies for the Prevention of Cruelty to Children, 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.[37]

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

[37] Lela B. Costin, Howard Jacob Karger, and David Stoesz, The Politics of Child Abuse in America 67-75, 84-101 (1996).

maintenance and social services for "dependent children." Until 1972, state welfare departments typically administered financial assistance and provided social services through the same workers. 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. A landmark was the 1962 survey in the Journal of American Medical Association of seventy-one hospitals that found 302 cases of non-accidental child abuse. The term "battered-child syndrome" was introduced into medical diagnosis.

States responded to such reports with laws encouraging and mandating reporting of abuse and neglect. The federal Child Abuse Prevention and Treatment Act of 1974[38] further encouraged and coordinated reporting. The statute 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 re-assessment. Investigations showed shocking administrative disarray in the state systems. Many simply could not account at all for large fractions of the 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.[39]

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

---

[38] 42 U.S.C. 5101 et seq.

[39] Marsha Garrison, "Child Welfare Decisionmaking: In Search of the Least Drastic Aternative, 75 Georgetown Law Journal 1745, 1756 fn. 44 (1987) (discussing a study finding, "[h]omemaker assistance as an option was discussed in about three percent of cases (as compared to 17% of parents who felt that a homemaker could have averted foster care) while daycare was discussed in less than two percent (as compared to 29% of parents who thought that daycare could have averted foster care").

care, adoption, and family support. The Act, as frequently amended, continues to provide the basic federal framework.

Substantively, the Act 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.[40]

Procedurally, the Act since 1980 has prescribed the kind of individuated attention that has become a central feature of current welfare programs. The case worker, 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. Goals are to be met through the provision of an array of services. The record must document the "appropriateness" of the services provided. The child's circumstances and the plan must be reviewed "periodically but no less frequently than every six months."[41]

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]

The dialectic of rules and standards has been a prominent theme in the AACWA. The 1980 text embraced family preservation through a standard. States were required to make "reasonable efforts" to avoid removing a child from his natural parents, and where removal was necessary, to reunify the family.[43] In 1997, Congress retreated somewhat from preservation. The use of foster care, which initially declined after

---

[40] Cite. The Act also creates preferences for less "restrictive (more family like)" placements, out-of-home placements that are close to where the natural parents live, and placements of siblings together. 42 USC 675(5)(a).

[41] 42 USC 671(a)(16), 675(1).

[42] 42 USC 5102a(a)., 5106(b)(2)(xi)

[43] 42 USC 671(a) (10).

1980, was trending upward again.[44] An important influence was the increase in serious abuse and neglect cases associated with the epidemic of crack cocaine. Another was the re-appraisal 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.[45]

Apparently fearing that the "reasonable efforts" standard was being treated as a virtually irrebutable presumption, Congress in the Adoption and Safe Families Act of 1997 (AFSA) 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, AFSA also mandated that states initiate termination of parental rights whenever a child has been in foster care for 15 of the most recent 22 months.[46] Thus, Congress initially prescribed a standard; then perceived that it was being treated like a rule strongly favoring reunification; and pulled back by prescribing an explicit rule limiting reunification efforts. Not surprisingly, the AFSA 22-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.

The dialectic of judicial and bureaucratic control is also prominent. The original legislation emphasized judicial modes of accountability. Congress mandated as early as 1974 that children have a lawyer "guardian ad litem" in any removal proceeding.[47] 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.[48] 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 12 months of removal of a child and at least annually thereafter. And removal of a child requires a

---

44 Marsha Garrison, "Reforming Child Protection: A Public Health Perspective," 12 Virginia Journal of Social Policy and Law 590, 591-92 (2005).

45 Id., at 607.

46 42 USC 675(5)(e).

47 42 USC 5106a(b)(2)(13); Pub.L. 93-247.

48 42 USC 671(a)(12).

judicial determination that the "reasonable efforts" at reunification required by the statute have been made.[49]

However, expectations for quasi-judicial and judicial oversight have been disappointed. Guardians ad litem 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.[50] 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 the "reasonable efforts" are being made in any permanency hearing, a 1989 study found that no such determination was made in forty-four percent of cases.[51]

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, where 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. Thus, family and delinquency court judges sometimes express frustration at the limits of their abilities to compensate for systemic problems and support for structural reform cases.[52]

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 re-orient such conventional monitoring in two

---

[49] 42 USC 672a(2)(A)(ii), 675(5)(B), 675(5)(E)(iii). The Act permits administrative rather than court hearings for these purposes if they satisfy certain conditions of openness and independence, but we are unaware of any state that uses this option. (ck)

[50] Utah example

[51] Guggenheim, cited in note , at 189 and works cited therein at note 18.

[52] For example, a juvenile court judge played an important role in the initiation of the Willie M. mental health case, which influenced the Utah-Alabama model discussed below. Kenneth Dodge et al, Willie M.: A Legacy of Legal, Social, and Policy Change 9-12 (Center for Child and Family Policy, Duke University, n.d.).

ways. First, the state (if it accepts certain federal support) must establish "citizen review panels." The panels, composed of volunteers, including some with expertise in children 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."[53]

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 to a "performance" orientation" in which the focus is on achievement of goals. In this spirit, the statute directs the Secretary 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".[54]

Yet, for the most part, these initiatives have not borne fruit. "Citizen review" panels have been formed and met only erratically. They often lack expertise and access to information. HHS's initial efforts at outcome-oriented review were fumbling and often arbitrary.[55] There are some indications that they have recently improved, but only after they have been reconceived along the lines pioneered in the efforts in Alabama and Utah that we discuss below.

In general, the federal legislative mandates have failed to coherently re-shape practice in the states. The progress we report below 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 be attached to natural family ties when parenting has been gravely

---

[53] 42 USC 671(a)(13).

[54] 42 USC 679b. 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, e.g., Carol Harlow, "Accountability, New Public Management, and the Problems of the Child Support Agency," 26 Journal of Law and Society 150 (1999). Exec. Order 12,866, 3 CFR 638 (1999), required that all federal agencies, when feasible, regulate by prescribing objectives, rather than specific behavior.

[55] See National Coalition for Child Protection Reform, "The Trouble with Child and Family Services Reviews," (Aug. 2003); at www.nccpr.org.

dysfunctional. And they continue to disagree over whether administrative dysfunction creates greater risks of inappropriate removal or of inappropriate failure to remove.[56]

B. Structural Judicial Intervention: 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 non-compliance 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.[57]

These lawsuits have been based on the substantive due process theory of Youngberg v. Romeo, on the AACWA,[58] on the federal Rehabilitation Act prohibition of discrimination in federally supported programs against the disabled,[59] and on state laws. Regardless of the particular substantive formulations of the claims, the remedies have tended to assume common patterns. For many years, the general tenor of the decrees has been 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.

A commentator noted in 1994 that "the common wisdom among litigators is that the decree should be as specific as possible."[60] Marcia Lowry, who brought several landmark structural reform cases, has asserted the underlying view with exceptional bluntness:

---

[56] Compare Elizabeth Bartholet, Nobody's Children: Abused Children, Foster Drift and the Adoption Alternative (1999) (asserting a systemic bias against removal) with Martin Guggenheim, What's Wrong With Children's Rights? (2005) (asserting a systemic bias in favor of removal).

[57] ABA Center on Children and the Law, cited in note

[58] Some were initiated prior to the Supreme Court decision in Suter v. Artist M., 503 U.S. 347 (1992), rejecting private enforcement of the statute. Some have been brought since 1994, when Congress partially overruled Suter. 42 U.S.C. 1320a-2. See Charlie H. V. Whitman, 83 F. Supp 2d 476, 483-84 (D.N.J. 2000). (discussing the extent to which the 1994 amendment overrules Suter.)

[59] 29 U.S.C. 794.

[60] Chris Hanson, "Making It Work: Implementation of Court Orders Requiring Restructuring of State Executive Branch Agencies," in S. Randall Humm, et al, Child, Parent, and State: Law and Policy Reader 224, 230 (1994).

> [P]eople 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.' [61]

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, where Lowry represented the plaintiffs.[62] 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 (12 hours initially 10 hours per year thereafter in general matters, 30 hours initially in prevention of abuse and neglect), caseworkers (one week initially and four days per year thereafter in general matters, 16 hours initially in prevention of abuse and neglect), and supervisors (one week initially and 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."[63] 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.

---

[61] Marcia Lowry, "Foster Care and Adoption Reform Litigation: Implementing the Adoption and Safe Families Act of 1997," 14 St. John's Journal of Legal Commentary 447, 453 (1999-2000).

[62] 564 F. Supp. 1030 (W.D.Mo. 1983). 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.

[63] Charlie and Nadine H. et al. v. McGreevey (D.N.J.), "Enforceable Elements of the New Jersey Child Welfare Plan July 19, 2004 –FINAL", at 2.

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 six months to two years. And it sets forth twenty-two basic systemic indicators (from "length of time in care for children with a goal of reunification" to "the recurrence of maltreatment for children who have been abused and neglected") and 98 "benchmarks" (from percentage of reports of child abuse or neglect as to which each child in the household has been interviewed within 24 hours to percentage of "new supervisors [who] receive the requisite pre-service training before supervising frontline staff") 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 re-oriented regime.[64] The need for re-direction 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."[65]

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

First, there is the danger of preoccupation with measurable and specifiable norms at the expense more amorphous but more important ones. 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 24 hours of custody, mandates that the child's health needs be "adequately met."[66] 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 quality of the medical care. As the New Jersey reform plan recognized, there is the danger that workers will come to see the decree as a "checklist" demanding formal

---

64 G.L. v. Stangler, 873 F.Supp. 252 (D.Mo. 1994)

65 NJ Period I Monitoring Report, at 7. In July 2006, the plaintiffs and the new Corzine administration thoroughly re-negotiated the decree. The next monitoring report found that "New Jersey is finally on a positive path," but stated that it was too early to expect measurable results.

66 564 F. Supp. At 1037.

compliance without concern for whether its underlying purposes are served. 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."[67]

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. At the frontline, command-and-control monitoring does not connect with the way social workers are taught to view their practice. 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.)

Another dimension of the problem is the tendency of the early decrees to define monitoring exclusively in terms of investigations and reports by outside officers and committees. Failure to internalize the goals and practices of monitoring reinforced the tendency of the system's insiders to see the reforms as oppressive intrusions.

Second, highly specified decrees freeze practice and inhibit adjustment in the light of experience. Modifications of a decree require the consent of the plaintiff or a court order, which normally requires a showing of changed circumstances.[68] Such a process does not contemplate routine

---

[67] Utah Division of Child and Family Services, The Milestone Performance Plan 8 (May 1999). Where relations between the parties are not cooperative, things can be much worse. See, for example, the complaint of the New York City agency, under a commissioner widely regarded as able and progressive:

> The adversarial nature of litigation … tends to freeze both sides into positions of mutual distrust. Even when the underlying issues are resolved, plaintiff's attorneys are reluctant to settle cases without provision for frequent monitoring reports and meetings for many years. Attorneys for the City are reluctant to re-open or modify decrees when circumstances change, for fear of opening the door to new issues plaintiffs' attorneys might raise.

New York City Administration for Children's Services, Protecting the Children of New York: A Plan of Action for the Administration for Children's Services (December 19, 1996), at

[68] See Rufo v. Inmates of Suffolk County Jail, 502 U.S. 367 (1992).

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.[69]

Third, command-and-control tends to produce decrees that are useless in setting priorities in the inevitable situation when there is noncompliance. This is 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 plan showed poor compliance or non-compliance 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."[70]

Fourth, 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 completed and filed has at best indirect value as either a measure of outcomes or a locator of failure (especially if the data do

---

[69] Utah Milestone Plan, at 8.

[70] NJ Pd I Monitoring Report, at 12. The panel's advice is consistent with contemporary business management lore. For example:

> Are 25 measures to many? Is it possible for any organization to focus on 25 separate things? The answer to both questions is NO! If a scorecard is viewed as 25 (or even 10) independent measures, it will be too complicated for an organization to absorb.
>
> The balanced scorecard should be viewed as the instrumentation to a single strategy. When the scorecard is viewed as the manifestation of one strategy, the number of measures on the scorecard becomes irrelevant, for the multiple measures on the Balanced Scorecard are linked together in a cause-and-effect network that describes the business unit's strategy.

Robert S. Kaplan and David P. Norton, The Balanced Scorecard: Translating Strategy Into Action 162 (1996)

not distinguish between a total failure to meet or file and an inadequate meeting or filing and are silent as to the specific nature and cause of the inadequacy). New Jersey's performance-based decree emphasizes important measures of outcome – for example, frequency of re-abuse in state custody, the length of time to permanency, the frequency of "re-entry" 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.

Some of the many "benchmarks" in the New Jersey scheme do have potential diagnostic value. For example, failure to adequately investigate and train foster parents may correlate with harm to children in foster care. However, New Jersey's initial scheme made it very difficult to make coherent use of such data because it lacked ideas about the relation of outcome and diagnostic indicators. An effective monitoring scheme needs a set of hypotheses that link outcome and diagnostic criteria. The hypotheses are necessarily tentative and revisable, but they give coherence to interpretation of the data and efforts to formulate remedial interventions.[71]

IV. The Emerging Model: Alabama and Utah

A. Introduction

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. Its origins lie in the mental health field, and it was first applied to children's protective services in Alabama. It has since been adopted to varying degrees in several other states. It appears to be most fully developed in Utah.

---

[71] According to Kaplan and Norton,

> A properly constructed scorecard should tell the story of the business unit's strategy through … a sequence of cause-and-effect relationships. The measurement system should take the relationships (hypotheses) among objectives (and measures) explicit so that they can be managed and validated.

They give as an example:

> If we increase employee training about products, then they will become more knowledgeable about the full range of products they can sell; if employees are more knowledgeable about products, then their sales and effectiveness will improve. If their sales and effectiveness improves, then the average margins of the products they sell will increase. Id.

Id. At 149.

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 re-think and re-structure 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 cases focused on Alabama's institutions for the mentally disabled. 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. [72]

After the R.C. reforms were underway, Vincent left government and established a consulting organization, the Child Welfare Policy and Practice Group. The group has worked on child welfare reform, often in connection with litigation, in about a dozen states. He and his colleagues sometimes collaborate with Groves, whose Florida-based organization, Human Systems and Outcomes consults with a range of public service programs in education and mental health, as well as child protection.

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 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 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."[73] There had been a "breakdown in the Corrective Action Process".[74] At this point, another visionary social worker -- Richard Anderson, then deputy director and later director of the state Division of

---

[72] The story of R.C. up to 1998 is told in Bazelon Center for Mental Health Law, Making Child Welfare Work: How the R.C. Lawsuit Forged New Partnerships to Protect Children and Sustain Families (1998) (hereafter "Bazelon")..

[73] David C. v. Leavitt, 13 F. Supp. 1206, 1212 (D. Utah 1998).

[74] Id. at 1207.

Child and Family Services – appeared and worked with plaintiffs' counsel to re-orient the process. Vincent became a monitor and consultant, with assistance from Human Systems and Outcomes.

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."[75] 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.

Although the Alabama and Utah decrees are powerfully innovative, some of their features can be found in most child welfare reform plans. Such plans virtually always involve commitments to "system investments," or infrastructure development, such as purchase of equipment and 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 15-20 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 through a desk audit of a sample of cases. The Utah process involves about 50 "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 30 days of the referral. If this case involved an allegation of medical neglect, did the worker obtain an assessment for a health care provider prior to case

---

[75] Utah Division of Child and Family Services, The Performance Milestone Plan 8 (May 1999). 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.

closure?[76] 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 are more distinctive -- first, a repudiation of rule-bound authority in favor of a contextual understanding of norms; second, a distinctive understanding of the relation between the administrative center and local units; and third, 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. By 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."[77] Norms have a rough hierarchy that permits prioritization. The system emphasizes to workers "the larger context of their decisions."[78] Exposition begins with a set of "core values" or "overarching principles".[79] More specific rules are typically understood as provisional. A few norms, mostly procedural and documentation ones, take a categorical form, but even they permit "documented exceptions," where the worker can explain and substantiate the reason for deviation. Utah styles its basic norms as "guidelines."[80]

---

[76] Utah Division of Child and Family Services, "Case Process Review Questions - 2004". Data tracking on reporting on some of these measures is federally required. See GAO, Most States Are Developing Statewide Information Systems, but Reliability of Child Welfare Data Could Be Improved, (GAO 03-809, July 2003).

[77] "Consent Decree," at 2, R.C. v. Walley, Civil No. 88-H-1170-N (M.D. Al.), June 5, 1991.

[78] Milestone Plan 9.

[79] These are basically the values expressed in the federal legislation: (1) child safety, health, educational progress, and psychological well-being; (2) preservation of families where consistent with 1, otherwise placement in the "least restrictive", most "family like" setting, with a preference of kinship over non-kinship placements, and for family over institutional placements; (3) stability and permanency; (4) preservation of a child's family and community ties even where out-of-home placement is required.

[80] Utah's written materials treat "regulations" or "laws" as side constraints. These are typically rules imposed by statute or federal regulation that typically take the form of procedural or documentation requirements or of basic values, such as prohibition of race discrimination, that are more concretely embedded in the guidelines. "Law" is thus implicitly identified with rigid prescription peripheral to the central creative tasks of the

Each section of the guidelines is preceded by a statement of the underlying "philosophy" of the norms.

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 “attachment” without any clear delineation of the practical implications of the theory. The more practical 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.[81] Training under the reforms focuses on a basic skill set, sometimes referred to as "the Practice Model", that includes “engagement” (interviewing and counseling), “teaming” (deliberation), ”assessment” (interviewing and counseling), “planning,” and “tracking” (follow-up), that are portrayed as central to every phase of casework. The 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 vindicating the principles. Frontline offices and workers have relatively broad discretion to apply the 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."[82] 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, then, as before the reform, limiting them to a menu of standard options provided by centralized procurement).[83] The central administration also

---

system. In fact, however, by moving the more creative practices to the center of enforcement efforts, the decrees in cases like Utah and Alabama are responding to a different conception of law.

[81] This was the focus of the Alabama curriculum prior to the consent decree reforms. Vincent telephone call. A recent GAO report complains that it is still the orientation of the training curriculum in D.C. GAO, D..C. Child and Family Service Agency 10 (GAO-06-1093, September 2006).

[82] Milestone Plan, 8

[83] Bazelon 26-60.

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 a 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 re-imagine 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. [84] In Utah, reform was begun at the same time across the state, but in each of the five regions, progress was evaluated separately, and a region could exit separately from active monitoring under the decree when it met certain milestones.

B. 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 decisionmaking, and most importantly, diagnostic monitoring. In the latter area, the reforms have produced a striking innovation – the Quality Service Review (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 in "the Practice Model". Others put greater weight on central facilitation of diagnostic monitoring, especially through the Quality Service Review. 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

---

[84] Bazelon , at

possible for administration to learn from local practice while correcting its mistakes.

1. 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. For example, if there is an ongoing problem of substance abuse that has not been identified or treated, then many of the threats to a child's safety, permanence, and well-being may continue long after the risk from the referring event is believed to be resolved."[85]

A training vignette illustrates assessment with a case involving the "Archuleta family" that begins with a report that young children have been left alone at night.[86] The parents eventually return home under the apparent influence of drugs. The children are placed with a maternal great aunt. The father, whose drug use constitutes a parole violation, is re-incarcerated. After extensive investigation, analysis identifies the following needs:

-- The children need a permanency plan. Returning home seems possible. Long-term care with the aunt seems an alternative.

-- Both parents have drug use problems.

-- The mother needs to increase her employment prospects so she can replace the income lost from the father's re-incarceration.

-- The mother feels socially isolated.

-- The parents need to maintain contact with each other after the father's incarceration.

-- The parents need greater understanding of child development and parenting skills.

-- One of the children has delayed speech.

-- Both children are anxious about the disruption of their family life

-- The 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 to consider underlying causes of problems, using available diagnostic tools and available data in a systematic manner. For example,

---

[85] Milestone Plan 9.

[86] U.S. Dept. 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 (visited ,2/1/2007).

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 done at school or to follow up on signals in the health records that the child had suffered oxygen deprivation at birth.[87] In another case that was discussed as exemplary at a recent Utah review (one from a randomly selected sample), the worker obtained a grant from a private charity for the nurse working on the case team and the foster mother to travel out-of-state to attend 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 mother later adopted the children.)

Assessment goes with "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. But the system of care idea requires a deeper individuation: "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…."[88]

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 day care and transportation for the children, an arrangement with the children's godparents to provide babysitting relief for the godmother, 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, English-as-a-second-language education for the mother.[89]

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.[90]

---

87 Northern qsr 2006, 12.

88 Alabama decree, par

89 HHS, supra

90 Bazelon 54. 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.

Customization can also involve logistical convenience. In a recent Utah case, when the worker and a father could not find a parenting class compatible with the father's work schedule, they obtained the course materials, and the worker tutored the father.[91]

A good plan must attempt to anticipate contingencies, and adjust in the light of experience. From the beginning, the worker 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 teen-ager who is expected to remain in foster care should include understanding and preparation for when he "ages out" of the system.[92] 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 teen-ager who had been put on medication for attention deficit disorder 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

---

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 grandmother. 66.

91 Northern Region report 2006, at

92

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.
Salt Lake City Region Report 2006, at 26.

crisis precipitated by an injury to the foster father, and to re-assess in the light of this contingency.[93]

By insisting on flexible and tailored responses to circumstances, the "practice model" strives to attenuate the contrast between the "preservation" and "rescue" approaches to child welfare.[94] It has 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 fall-back options in the event that proves impossible.[95] Even where there is no prospect of reunification, continued parental contact often proves possible and desirable.[96]

One of the cases we reviewed involved a highly successful long-term foster care placement in which the foster parents and the birth mother had a cordial and collaborative relation and there was frequent contact between mother and child. The mother, who agreed that she was unable to act as primary caretaker, was an active participant in the process. She assisted the foster parents in getting the child to take his medication and immunizations, which he would not do for anyone else.

2. 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

---

93 Case reviewed by Simon in Northern Region QSR, January 2007. See also:

> At the time of review, the transitions were completely unacceptable. Not only has the next age appropriate transition not been managed, the actual transition to the … boys group home has moved (the youth) into the exact opposite direction away from individualized permanency. The transition from the foster home to the group home in to the south residential care facility to the boys open residential home were all done within a few weeks. (The youth) didn't understand why he had to be in all of those programs again as he thought he had successfully graduated from them years ago. This appeared to have left him with a profound feeling of failure and sense of futility. The staff at the boy's residential facility the reviewers visited with had not even talked with (the youth) yet about his transfer to the unit, leaving the reviewers with the impression that no effective transition planning had occurred between the [group home] South and West residential units to help (the youth) with the move. We do not know how effectively the other transitions were handled, and saw no documentations in the records that had been done.

Salt Lake City Region report 2006, at 30.

94 See notes above.

95 Utah guidelines 300. __

96 See Bazelon 39

out-of-home care caregiver and community resources."[97] The child if over 10 (Alabama) or 12 (Utah) is regarded as a "partner" in his or her plan.[98] So 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 right-holders and as sources of information and support. Involving parents and giving them a sense of "ownership" in the plan is an important worker duty. Review scores for "system performance" are reduced when parent participation is unsatisfactory for whatever reason. If a parent cannot make a meeting during work hours, the caseworker is expected to schedule one outside of work hours; if a parent cannot easily travel to the agency, the caseworker is supposed to schedule the meeting at his or her home or some other convenient location.[99] Children in foster care are expected to have regular contact and visitation with their natural parents, except in extreme situations.[100]

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 case worker -- 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. 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

---

97 Utah Guidelines 300.4(C).

98 Id. at 10.1; Alabama decree, par 44.

99 Family participation was considered unacceptable in six of the 24 cases reviewed in the Utah Northern region in 2006. A typical criticism was:

> Family members are not participating in team meetings. Meetings are held during the time that Mother's brother and his wife work; otherwise they would have been able to attend. Her grandparents and father have helped [the child] in many ways; yet they are not participating in team meetings. The [guardian ad litem], with whom the mother has a contentious relationship, does attend team meetings and as reported by other persons interviewed, is a major influence in directing the course of the case plan. Mother feels that she has little input into her case plan. She feels that "they put it together and I do it." 17

100 Alabama decree, par ; Utah guideline

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 important duty of the worker is to assemble the team, convene it periodically, and make sure its members are informed of facts, and are engaged in the planning.[101]

An example of the kind of re-framing such collaboration is designed to encourage occurred in a de-briefing review session one of us participated in. The case involved an autistic teen-age 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 in school, and his care taker 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 re-evaluated, less positively.

3. 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.[102]

---

[101] The trajectory of reform can be seen by comparing the team provision of the first G.L. decree, which mentions five categories of team participants, 564 F. Supp. at ; with the corresponding provision of the revised decree, which mentions 15 categories of participants. 873 F. Supp. at 257.

[102] The Utah plan, which has an exceptionally sophisticated monitoring regime, also provides for the following forms of assessment:

-- a survey of agency staff on their views of the progress of implementation

-- review of "trend indicators" on "annual outcomes" (e.g., "percent of children with substantiated allegations of abuse while placed in out-of-home care," percent

However, the reformers in Alabama use a distinctive monitoring procedure designed to complement the customizing and collaborative features of their practice model. The Quality Service Review (QSR) was first applied in child welfare in Alabama and has since been applied to child welfare programs in 11 other states, including Utah. In both states, the QSR became the central measure of compliance for purposes of decisions to terminate the decree.[103]

According to Ivor Groves, the "fundamental assumption" of the QSR is that "each case is a unique and valid test of the system." 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. Groves contrasts the "qualitative" QSR approach to the "audit focus" of conventional monitoring:

> AUDIT FOCUS: Is there a current service plan in the file?
>
> QUALITATIVE FOCUS: Is there service plan relevant to the needs and goals, and coherent in the selection and assembly of strategies, supports, services and timelines offered?[104]

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 72 cases for Salt Lake City (out of a total population of about 1,500 cases), and 24 for the four less populated regions. In Alabama's county-run system, samples range from 68 for the largest county and 12 for the smaller ones. Random 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, older and younger children, and boys and girls.

The cases are reviewed by teams of two. Ideally (but not always in practice) 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 of the

---

of children who re-enter out-of-home care within six months," "average length of stay of cohorts of children in out-of-home care," "number and percent of children who attain permanency."

-- survey of views of foster parents

-- report on staff turnover

Milestone Plan 22, Appendix.

103 On the centrality of the QSR to Alabama and Utah, see TAN notes below.

104 Cite

two 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, non-family 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 individuated 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 30 days. The other indicators concern "system performance" over the past 90 days. The two sets of indicators respond to the goals of what business scholars Robert Kaplan and David Norton 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.[105]

The principal QSR instrument is a 95 page booklet with a series of suggestive questions for each indicator. (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?")

Scoring for individual indicators is on a six point scale with 6 "optimal" and 1 the worst. For aggregation purposes, some indicators

---

[105] Cited in note , at 150.

receive more weight than others. On Utah's instrument, for example, "safety" and "permanence" are multiplied by three; "learning progress" by two; and client "satisfaction" by one in determining "overall status". "Planning" and "assessment" are multiplied by three; "team coordination" by two; and "caregiver support" by one in determining "overall system performance." A score of four on a given indicator implies that status or performance on the relevant dimension is "minimally acceptable", and an average score of four or higher indicates that overall child status or system performance 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 matrix such as this:

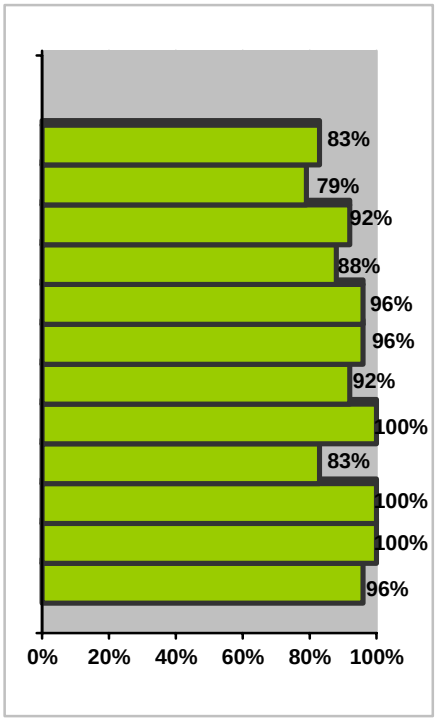
Child Status and System Performance Ratings			
Child Status	Rating	System Performance	Rating
1a. Safety of the child	5	[redacted]	
1b. Safety Risk to Others	5		
1. Overall Safety	5		
2. Stability	5		
3. Appropriateness of Placement	6		
4. Prospect of Permanence	4	1. Child and family Participation	5
5.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

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

When cases are aggregated, a matrix for the system or subsystem looks like this:

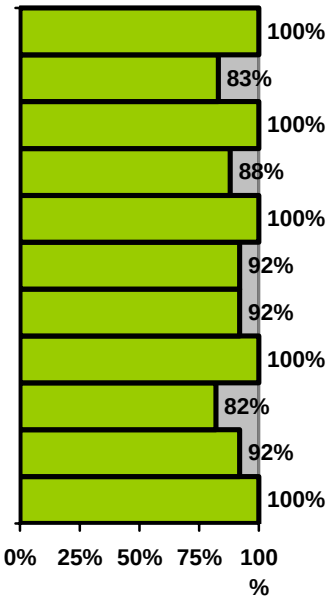
Overall System

Northern System Performance									
	# of cases (+)	# of cases (-)	Exit Criteria 70% on **Shaded** indicators/ Exit Criteria 80% on overall score	FY03	FY04	FY05	FY06	FY07 Current Scores
C&F Team/Coordination	20	4		42%	67%	75%	71%	83%
Functional Assessment	19	5		42%	54%	60%	54%	79%
Long-term View	22	2		25%	58%	71%	75%	92%
C&F Planning Process	21	3		46%	63%	79%	83%	88%
Plan Implementation	23	1		71%	71%	83%	88%	96%
Tracking & Adaptation	23	1		67%	71%	88%	83%	96%
C&F Participation	22	2		50%	88%	96%	67%	92%
Formal/Informal Supports	24	0		75%	79%	96%	92%	100%
Successful Transitions	19	4		63%	73%	83%	82%	83%
Effective Results	24	0		75%	71%	96%	92%	100%
Caregiver Support	13	0		94%	92%	92%	92%	100%
Overall Score	23	1		58%	79%	83%	88%	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%	96%	96%	100%
Stability	20	4		79%	75%	92%	75%	83%
Approp. of Placement	24	0		100%	96%	96%	100%	100%
Prospects for Permanence	21	3		42%	67%	71%	71%	88%
Health/Physical Well-being	24	0		100%	100%	100%	100%	100%
Em./Beh. Well-being	22	2		88%	79%	75%	92%	92%
Learning Progress	22	2		79%	75%	83%	92%	92%
Caregiver Functioning	14	0		88%	100%	100%	100%	100%
FamilyResourcefulness	14	3		44%	56%	76%	71%	82%
Satisfaction	22	2		75%	92%	100%	96%	92%
Overall Score	24	0		100%	100%	96%	96%	100%



The reviewers discuss among themselves the indicators and information from their cases that might explain their significance. They meet collectively with personnel from the region to discuss the systemic significance 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, which together transform traditional clinical practice. First, it's a form of clinical training for the caseworkers and their supervisors, an "action-oriented learning process," as the designers describe it.[1] 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.[2] 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."

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 them through analysis of cases. The scoring system forces the reviewers to formulate their judgments in forms 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

---

[1] Human Systems and Outcomes, Quality Service Review: Protocol for Use by Certified Reviewers 2 (Wisconsin version 2005)

[2] See Carlton Munson, Supervision in Social Work: Classic Statements and Theoretical Issues (1979).

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.[3]

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 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,[4] it is rarely plausible to get a large enough sample size for statistical validity. 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 workers 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" and its "zero tolerance" for major defects.[5] 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

---

[3] The U.S. Administration for Children and Family Services initially took the contrary position for purposes of its performance reviews, but it has since adopted the Utah view. Cite

[4] Brad McGarry, director of Utah's Office of Services Review, estimates that the costs of QSR review average about $3,000 per case.

[5] See Philip B. Crosby, Quality Without Tears 74-84 (1984).

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 in-home or protective service cases and foster care cases. Indications that children in one group are faring worse than another are treated as signals to re-assess 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 reason for several years. The region focused training efforts on perceived problems in these cases, and the difference disappeared in the 2006 review.[6]

C. Systemic Assessment in Public Law Litigation: Termination in Alabama and Utah

---

[6] Western Region Report 35-36 (Dec 12, 2006). 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. Telephone interview with Linda Winneger, Director of Milestone Planning, Utah DCFS, March 19, 2007.

Another example: Utah analysts found statewide that status scores in a major sub-category of foster care cases -- where the children are placed with relatives -- were lagging other sub-categories. Administrators responded by focusing additional review efforts on kinship placement cases. They concluded that decisionmakers 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 more general experience and experience with their own kids. Training efforts were re-configured 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. Telephone interview with Richard Anderson (Aug. 18, 2006).

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 and more impressionistic data to assess performance.[7] After finding that the state had made the required "system investments", in personnel, training, and information technology, the court proceeded to consider the quality of care. It defined compliance primarily in terms of the benchmark, agreed to by the parties, of 85 percent of cases scoring at least acceptable under a QSR review supervised by the monitor. Each of the state's 67 counties had met the benchmark at least once in recent years, but some had subsequently regressed below the benchmark. In 2005, the court had directed the monitor to do a further review. The monitor reviewed 152 cases in ten representative counties. Seven counties passed, but three did not.

Turning to the more specific goals of the decree, the court found that the state had achieved adequate performance on child safety, relying on QSR "status" scores of 90 to hundred percent. It also found that the timeliness of investigations of abuse-and-neglect complaints met was within the benchmark norm, relying on case processing data.

The court also saw favorable trends in federally-prescribed reporting indicators: The median time a child remains in state care had been falling. "Re-admissions to care" -- cases that have to be re-opened because dispositions prove unsuccessful -- had declined. The number of adoptions had increased. The average number of placements experienced by children in care had declined by a full placement since the inception of the case.

In other areas, the court found failure to meet goals. The remarkable achievements in family preservation in the early years of the decree[8] had partially reversed in the years after 1995, when out-of-home placements increased. Caseloads and staff turnover sometimes exceeded standards. Despite dramatic improvements in QSR-measured assessment practice, there were some counties in which plans "were not driving practice" for a substantial portion of the caseload. There was extensive failure to meet the

---

[7] R.C. v. Walley, 2007 WL 118101 (M.D. Ala. 2007).

[8] In the first three years, out-of-home placements dropped 20 percent; no other state experienced any drop during this period. Bazelon 57.

requirement of the 1997 federal legislation that termination of parental rights be initiated after 15 months in care.

The court declined to treat the federal 15 month deadline as categorical for the purposes of the decree, accepting the monitor's judgment that such a "rigid time guideline" is neither plausibly attainable nor in the interests of many children.[9] Instead, the court looked to the individualized judgments of the QSR. Permanency was considered acceptable in 83 percent of the cases in the sustaining counties and 63 in the non-sustaining ones. Moreover, the court refused to assess permanency in isolation from the adequacy of placement. (An isolated focus on permanency encourages a rush to get children into placements that are long-term but not necessarily optimal.) QSR scores on placement were at least adequate in between 90 and 100 percent of cases in all the sample counties.

With respect to the other shortcomings, the court emphasized that, in each case, the state had, through established processes, diagnosed the problem and initiated plausible corrective action on its own. 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, which data showed, accounted for as much as 70 percent of intakes in some counties. It noted that the defendant had inaugurated a collaboration with law enforcement to deal with such cases. With respect to caseloads, statewide and county staffing committees tracked caseloads closely and re-assigned workers in response to overloads. The state had recently increased pay in response to turnover. It had responded to deficiencies in planning with increased training and an effort to simplify the planning process.

The court's analysis concluded by briefly noting that the department planned to retain much of the monitoring regime constructed under the consent decree, including its "quality assurance" regime. It noted that the regime issues "report cards" published on its web 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.[10] Its discussion of how quality assurance

---

[9] At 44-45.

[10] 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."

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, and none of 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 to 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 commits to a final round of review by the monitor to determine, among other things, that "the integrity of the [QSR] process has been maintained." It further commits to maintain until at least the end of 2010 the key elements of what we have called the Alabama-Utah model in place. 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."[11]

With respect to external accountability, the agreement contains several provisions requiring that performance data, including QSR data, and reports be published on-line. It requires various periodic reports to the public, the legislature, and non-governmental organizations. The agreement requires increased support for statewide and regional Quality Improvement

---

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 a third of workers had excessive caseloads. Nearly a 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 13 months or longer had increased from 46% to 65%. "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., at 79-89.

[11] "Agreement to Terminate Lawsuit," David C. v. Huntsman (Civil No. 2:93-CV-00206TC, May 11, 2007) (D. Utah). Par.s 18-30, Appendix B.

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.[12]

V. 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.

A. Federal Oversight

In the experimentalist view of American democracy, the role of the federal government in relation to the states parallels the role within states of state to local administration. Federal administration plays an oversight and facilitation role. It holds 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.[13]

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 of the Department of Health and Human Services (ACF) has developed an assessment process -- the Child and Family Services Review (CSFR) -- that shows great promise. Despite important 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,[14] has evolved toward increasing resemblance to the Alabama-Utah model (in part,

---

12 Id., par.s 31-35. On the composition of the committees, see Milestone Plan, at 78.

13 See Dorf and Sabel, cited in note , at 419-37.

14 Some such process is mandated by 1994 legislation. 42 U.S.C. 1320a-2a.

under the influence of Alabama[15]). 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.[16]

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.[17]

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.[18] Yet, the flaws seem remediable, and the trend of federal practice is promising.

B. 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.

---

[15] See Center for the Study of Social Policy, "Improving the Performance and Outcomes of Child Welfare Through State Improvement Plans (PIPs)," May 2003, pp. 28-29 ("Much of the basic protocol for the federal CFSR being deployed by the U.S. Children's Bureau grew from seeds planted first in Alabama.")

[16] Comments on the Second Round: An Interview with Jerry Milner, Senior Child Welfare Specialist, *Child Welfare Matters*, Spring, 2007

[17] Pennsylvania Office of Children, Youth & Families, Program Improvement Plan Status Report Final Report - May 2, 2005**:** Presented to Administration for Children and Families; State of Wisconsin Department of Health and Family Services, Division of Children and Family Services: DCFS Memo Series 2005 - 08 /Action September 14, 2005 Re: Continuous Quality Improvement (CQI) Program.

[18] Child Welfare Policy and Practice Group, "Comparison of Case Review Tools: Child and Family Services Review (CFSR) and Quality Service Reviews (QSRs),"(2005) (on file with authors); GAO report 2004.

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. The first system will have more children who start out in relatively good shape (including some for whom intervention was unnecessary).

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, 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.[19] You cannot measure the efficacy of a

[19] For a discussion of methodology and an illustrative study, see Michael S. Wald, J.M. Carlsmith, and P.H. Leiderman, Protecting Abused and Neglected Children (1988).

Abhijit Banerjee has 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. "Inside the Machine: Toward a New Development Economics," Boston Review (March/April 2007), at 12. Education is one of many aeas in which QSR-type processes could play an important role in assessment and improvement of complex services.

program unless you can see the extent to which it is actually being implemented. The QSR is the best means of doing so.

VI. The Antinomies of Welfare Rights Revisited

We now reconsider the basic tensions of welfare rights in the light of the Utah-Alabama model of child welfare program assessment. In each instance, we find that the model has jurisprudential qualities that elude the dichotomies that dominate prior discussion.

A. Rules v. Standards

The Alabama-Utah model involves a distinctive resolution of the rules/standards dialectic.

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 decisionmakers 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. Decisionmakers do not expect deference to ineffable wisdom attested to by professional credentials. Their assessments and planning have to be explicit and written. If virtually every rule is subject to implicit and explicit exceptions, these exceptions must generally be "documented exceptions."[20] Good case work is not considered adequate if it is not consistent with, and understood by the practitioners to derive 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. It makes it easier for new members of the team to acquire understanding of the team's prior work. Equally importantly, 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.

---

[20] See Utah manual

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.[21]

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 policymakers, 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

---

[21] Child Welfare Policy and Practice Group, "Implementation of Alabama's R.C. Consent Decree: Creating a New Culture of Practice" at 5, n.d.

"best interests" of the child are without relying on any background theoretical or ideological premise.

For example, in the first few years of the new regime in Alabama, decisionmakers 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." The QSR system identified these cases, and administrators responded with corrective discussion, training, and guidelines.[22] This re-orientation occurred before Congress's 1997 re-formulation of the AACWA reunification presumption to reduce what the legislature then perceived as excessive commitment to re-unification. In a system with the capacity for careful contextual response and adaptation, the presumption was relatively unimportant.

B. Bureaucratic v. Adjudicatory Internal Control

The motto, "Every case is a unique and valid test of the system," expresses the QSR'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've 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 decisionmakers 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, the decisionmakers reviewed, 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

[22] Vincent interview.

deliberation. 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 directly 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. All systems have some form of stakeholder-initiated administrative review, such as a hearing procedure or an ombudsman. Such a process is clearly an important constituent of a welfare system. 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 first better if they were relieved of much of the burden of routine supervision they must now perform where administrative systems malfunction severely.

C. Discrete v. Systemic Judicial Intervention

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 responsibility 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. At the same time, it is a

fundamentally systemic intervention that seeks to change the system as a whole.

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."[23]

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 holding of Suter[24] and the cases that preceded it denying a private right to enforce the 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, 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. 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.[25] Suter's effort to narrow intervention would have been counter-productively disruptive had it not been ineffectual. Confining the court to narrow intervention constrains its discretion only at the cost of arbitrariness.

D. Negative and Positive Rights

---

23 Milestone Plan, at 5.

24 Cite, see note above.

25 Bazelon , at 16-17. ;

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 the state's failure to provide services to troubled families increases the likelihood that children will be removed. It would be senseless and incoherent to give priority to a negative right against unnecessary child removal over a positive right to services that would obviate removal.

At the same time, 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.[26]

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 the relation of the claimant's interests to the relevant public purpose, qualified decisionmakers, and at least minimal participation by the claimant. "Accountability" implies a reasoned explanation by the decisionmakers, 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.

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.[27]

The shadow of such a conception of social rights can be found in the holding of the Supreme Court in Youngberg v. Romeo that due process requires certain welfare decisions to be "based on professional judgment." Youngberg is impaired by its limitation to rights that can be styled as "negative" (for example, on the facts of the case, conditions of state

---

[26] Charles F. Sabel and William H. Simon, "Destabilization Rights: How Public Law Litigation Succeeds," 117 Harvard Law Review 1015 (2004).

[27] For elaboration, see id.

custody) 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've 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[28] and Minister of Health v. Treatment Action Campaign (TAC)[29]. 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 anti-retroviral 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

---

[28] 2000 (11) BCLR 1169 (CC) (S. Afr.)

[29] 2002 (5) SALR 721 (CC) (S.Afr.)

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 re-allocation 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."[30]

Moreover, states typically find that good management increases access to outside funding. A key component of most reform plans is the re-organization of efforts to maximize federal re-imbursement 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."[31] No doubt it is an exaggeration, but it expresses the basic insight that value can be created through re-organization.

---

[30] 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.

[31] Crosby, cited in note , at

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, transparency that the court imposes on the defendant can be used to assess the legitimacy and efficacy of the court's 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 further distinctive innovation in the possibilities of legislative oversight. More than a dozen legislators have participated in the QSR process as reviewers.[32]

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 anti-welfare 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 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 morale support and information exchange, but it also played an important role in lobbying the legislature for adequate resources.[33] 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

---

[32] Interview with Richard Anderson. 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. Matthew McCubbins and Thomas Schwartz, "Congressional Oversight Overlooked: Police Patrols versus Fire Alarms," 28 American Journal of Political Science 165 (1984). The QSR suggests an alternative that could provide far richer information.

[33] Bazelon, at 76, 9-83.

vulnerable stakeholder constituencies in ways that make the reforms politically more stable.