

"Unintended consequences of the push for accountability: the case of national child welfare performance standards," by Mark Courtney, Barbara Needell and Fred Wulczyn, Children and Youth Services Review, December 2004 is an exasperated challenge of the performance standards being utilized by the federal government in the child and family service reviews (CFSRs). Fred Wulczyn and a number of other academic experts regarding out of home care have argued for several years that states and the federal government should be utilizing entry cohorts of children instead of point in time or exit cohort data to describe and evaluate out of home care programs. These scholars have repeatedly made the point that counts of children in out of home care at a point in time, e.g., on any one day or during a three month period of time, have the inherent weakness of excluding children who have entered and already exited foster care, thus creating a biased look at foster care / kinship care which exaggerates length of stay. "Cross - sectional samples (e.g., all children in care at a particular point in time) tend to be biased toward those children who have relatively long stays."

"In contrast, samples from exit cohorts (children who have left the foster care system) tend to be biased toward those children from recent entry cohorts with relatively short stays in care since those with longer stays have not yet exited care." The CSFR permanent planning performance standards are based on data from exit cohorts. States' permanent planning data from exit cohorts is likely to mask the extent to which large percentages of children entering out of home care are experiencing long lengths of stay and fail to accurately identify states with much lower median lengths of stay than average. In doing so, the sampling bias created by use of exit care data "makes it impossible for entities whose performance is being assessed, in this case states, to accurately measure their success in improving performance over time." "Second, (sampling) bias creates an unfair basis for the federal government to determine which states must embark upon performance improvement plans and upon which states to assess financial penalties."

These researchers and policy analysts insist that "state performance should be evaluated using data that best reflect the experiences of all the children served." "For analyzing performance, it is in most instances best to group children who entered out - of home care during a specific year, following their progress through placement until they experience some form of exit." "Because entry cohorts include all children (in out of home care), the inherent biases are far more limited and the data are more easily interpreted, especially if officials are trying to track change over time."

These researchers go on to compare six states' (from the Chapin Hall Center Multistate Foster Care Data Archive) entry cohort and exit cohort data for reunification and adoption. As expected, they find that "exit cohort data significantly overestimate both the proportion of children who are reunified within 12 months and the proportion of children

adopted within 24 months." "This is completely consistent with the idea that exit cohorts favor children who have been in foster care for shorter periods." "This same pattern characterizes the adoption data."

"Second, performance trends -- changes in outcomes over time -- are sensitive to whether one considers only children who are exiting care or if one chooses to look at a series of entry cohorts." "In fact, if the experience of these six states is any indication, one should expect the trends observed for exit cohorts to differ from those observed for entry cohorts rather than to be similar." In short, tracking states' trends over time using CFSR performance standards is likely to produce highly distorted accounts of how states are doing in child welfare reform, according to these researchers. "This is an environment that is ripe for quick - fix consultants that have nothing worthwhile to sell and administrative gaming strategies that "improve" measured outcomes at the cost of serious and unintended consequences."

These authors also state that "Even putting aside the fact that current federal measures are likely to lead the states astray, the knowledge base regarding the relationship between population characteristics, programs and policies, and child welfare outcomes is so poor as to offer little sound guidance for program improvement." "A better approach would be for the federal government to see the movement towards outcome measurement as an opportunity for the field of child welfare services to learn about the relationship between policy, program, and outcome." "... the primary focus of the Child and Family Service Review process should be to better understand why states are achieving particular constellations of outcomes," rather than a process for bringing states into "substantial conformity" with federal standards.

Courtney, Needell and Wulczyn are among the most prestigious child welfare scholars in this country. Their views have been heard (due to their status) but apparently disregarded, thus the exasperated and even acid tone of the concluding paragraphs. "Clearly, a federal bureaucracy that is unable or unwilling to improve on the flawed performance measures that are now in place is ill equipped to move the field forward." "Our experience tells us that very few states are in a position to do much better."

"In the past 10 years, states have vastly improved their data systems, but have not done commensurate work in developing the human resources necessary to make use of all the new data." "States should replicate, perhaps with the help of the federal government, promising partnerships between state child welfare agencies and university - based researchers." "This capacity will not appear overnight, but it will not appear at all without a serious and sustained commitment of resources."

Practitioners may wonder why these authors are so agitated about what seem technical and abstruse research issues; but part of the reason for agitation is the understanding of researchers that one can damage programs rather than help them by measuring the wrong things and pretending to understand relationships between policy, practice and outcomes that have not yet been established by sound research.