The opportunity and challenge of collecting street-level information

With the hope of changing the lives of youth in foster care, in 1999 the U.S. Congress enacted the Foster Care Independence Act, also known as the Chaffee Independent Living Act (Act). The Act provided $140 million in block grants to states to support youths' transitions to independent living and required the Federal Administration for Children and Families (ACF) to develop a national data collection and reporting system. The system, known as the National Youth in Transition Database (NYTD), seeks to track outcomes of youth receiving independent living services and to trace certain youth outcomes over time, even as they age out of the foster care system.

The NYTD follows a long history of federally mandated reporting systems. However, it is the first national, longitudinal data collection effort focused on gathering data directly from the youths themselves, and not just defaulting to administrative data as the source of evidence. To get this type of first-hand data directly from youth requires not just policy at the federal level, but also action and resources from state and county agencies, as well as the youth. From our experience, the cost to government agencies that do not possess the technical capabilities required to develop innovative data collection, transmission, and analytics make it increasingly difficult to find value in this data and to be compliant with reporting.

In 2010, the Center for Technology in Government at the University at Albany (CTG), partnered with the New York State Office of Children and Family Services (OCFS) to design a model for data collection for NYTD that fit OCFS' current technological and structural environment. While the experiences in the first wave of NYTD data collection have varied, the experience in New York sheds light on ways to improve upon the NYTD design and system going forward, as well as provide similar insight for efforts in other policy domains.

NEW YORK'S STRATEGY

The New York data collection model included phone and paper surveys. CTG recruited another partner, the Center for Survey Research at Stony Brook University to conduct the telephone surveys. Paper surveys were coordinated by CTG, but administered to youth through caseworkers at facilities. To acquire the completed surveys, the agency team, CTG, and survey staff had to solve a mix of interrelated problems that grew primarily out of the complex environment of the foster care system.

NYTD AS "CLOCKWORK" REPORTING On February 26, 2008, the final data collection design emerged in the form of a final rule (Federal Register 73 FR 10338) that established the National Youth in Transition Database (NYTD). ACF and its various collaborators took almost ten years to decide on the types of information needed to fulfill the intention of the legislation and to develop a plan for data collection.

The ACF designed and mandated a data collection and reporting program that had a clockwork-like logic: each state extracts data from the local agencies, providers, and youth; then this data flows back to Washington, then the data is disseminated by the federal government and used for various analyses. This logic is described in more

detail below:

1. States use foster care data to identify all youth turning 17 while in care and identify their location.
2. Surveyors (whether hired or in-house) locate and offer the questionnaire to all youth within 45 days of their 17th birthday.
3. Youth agree to participate and complete the survey.
4. Surveyors re-locate the same youths that completed the survey at 17 again at two-year intervals (at 19 and 21).
5. The 19 and 21 year olds, whether in or out of care, are invited, agree to and complete the follow-up surveys.
6. The state agency stores results in an NYTD compliant database and transmits the data to the ACF on time.
7. Repeat baseline cohort (a new group of 17 year olds) at three year intervals.

What this simple logic fails to take into account is the administrative complexity of the foster care context and the willingness of youth to participate. The foster care system supports hundreds of separate organizations, distributed data collection processes, extensive possibilities for data errors, and inevitable gaps in the youths' records.

NYTD DATA COLLECTION CHALLENGES The challenge of the NYTD data collection effort is that it reaches all the way to individual youth, and requires tracking some of them for five years during which many, if not most, have left any formal state foster care program. The NYTD is designed to collect administrative data from state agencies and additional data directly from youth in the transition population through surveys. Starting in October 2010, states were required to collect data on cohorts of foster youth, beginning with those who reached their 17th birthday between October 1, 2010 and September 30, 2011.

Furthermore, state supervised, locally administered systems like New York's have few clockwork qualities. Foster care in New York involves 57 semi-autonomous counties outside of New York City and hundreds of voluntary service providers of all sizes. And, like other states, and perhaps most importantly, New York faces the significant challenge of maintaining contact with youth over several years and securing participation with a population of mobile, diverse, and often circumstance-challenged youth.

The main problems resulting from clockwork type assumptions for data collection, as implemented in a complex foster care system, are outlined briefly below.

Direct participation. Administrative data is collected on individuals receiving services as a matter of course without those individuals necessarily knowing it. It is standard practice and allowed, mainly because that data is stripped of identifying information. The NYTD requires youth turning 17 to actually complete the survey themselves. Youth in this cohort are minors and the state, local providers, caseworkers, and parents have a responsibility to protect the privacy and safety of the youth. Despite outreach and education about the merits and legitimacy of the NYTD, various gatekeepers to the youth refused access.

Contact information accuracy. Casework systems are designed to serve several purposes. Therefore, the fitness for use of the state's administrative databases or local databases was not as helpful as one might anticipate. In a complex administrative system like New York's, the data elements related to "contact information" may have many different meanings depending on uses and who owns the data. The contact information in the system may be accurate and correct for billing purposes or for other legal requirements, but is not set up to track movements of individual youths within an agency. Contact information may be in the "notes" sections of the database systems or in the paper files of caseworkers. The data quality issues we experienced with contact information vividly illustrates the extent of the problem: Approximately 80 percent of the initial contact information we received from the state systems was not fit for our purpose or incomplete or both. That does not mean it is not fit for other purposes already established and required by previous mandates for data collection.

"It is crucial that federal, state, and local governments work together to develop new tools and systems that facilitate better practices and help improve outcomes. The goal is to collect and report data that is useable to help policymakers, program administrators, and caregivers at all levels to provide better services and support to our youth." -*Commissioner Gladys Carrión, Esq., Office of Children and Family Services in New York State*

NEAR REAL-TIME REPORTING While not designed as a real-time data collection and reporting model, the NYTD approach mandates certain reporting time frames. The first timeframe is that youth must take the survey within 45 days of their 17th birthday. The second is that states must transmit data back to Washington every six months during the reporting period. In our experience, the average time to complete a survey was around 30 days, mostly due to the need to navigate around local gatekeepers. When contacting over 100 different agencies,

the surveyors had to navigate different organizational structures and work practices to find the right person that could put them in touch with the youth. State and county records were of little help in this regard. Even with cooperative respondents, the many organizational layers often slowed responses and prevented contact during the 45 day window. Every delay due to additional time needed to repeatedly explain participation or correct contact information errors reduced the likelihood of contacting the youth within the required time period. This was a symptom of the difficulty, in spite of repeated efforts, of establishing effective communication lines from the state level down to the hundreds of different locations and care givers spread throughout the state, and maintaining that relationship.

When data collection requirements and guidelines are designed from afar, the distance between that view and the reality on the ground results in a number of unexpected consequences. These include overly cumbersome data collection processes, less than adequate data, and mismatches between data collected and data needed, among others.

IMPACT OF COLLECTION ON STATES AND COUNTIES The legal responsibility for the NYTD falls on state agencies, but for a national project like this to succeed, states together with counties must bear the main data collection and reporting burden, both the benefits and the costs.

The weight of the data collection and reporting burden can vary substantially from state to state and county to county depending on their foster care systems and technical and administrative capabilities. In October 2010, some states already had information and management systems designed for and capable of responding to most of the NYTD administrative data reporting requirements. And any state with such an existing information system and administrative authority for foster care would be well positioned to find and survey the youth as they reached their 17th birthday. Other states, those with data in multiple statewide and local level systems, were not configured to respond in the same way to the NYTD task.

For state-run, county-administered settings, a large portion of the administrative authority for foster care resides at the county level. These states and county agencies faced different, and in some cases, a more difficult challenge in becoming NYTD compliant. Non-compliance risked financial penalties to states, penalties that could reduce foster care funds for counties as well.

While the costs of data collection are shared across state and local levels, so is the benefit. Improved data about foster youth in transition can help policy makers, program administrators, and care givers at all levels to provide better services and support. These efforts are sure to be a difficult and costly undertaking, involving individual service providers and supervisors in foster homes or voluntary agencies, county social workers and supervisors, and state agency IT and program operations.

LOOKING TO THE FUTURE Our experience implementing the survey portion of NYTD brought to light very clear and important issues regarding developing new types of data resources. Policy makers and practitioners should be looking at future information supply issues. Is the future in clockwork data collection systems like the NYTD or in opening data? Can open data strategies solve the need to create new data resources?

Opening government data may not be the answer to all the lack of information problems government faces. So much of what government does, especially in the areas of complex service delivery, such as welfare, education, or child protection, or relating to accountability or transparency is not collected through agency legacy systems. The NYTD clearly shows that the needed information tends not to be part of government agency legacy systems. Legacy systems house governments' machine-readable data, and the data within them are often collected for specific purposes and programs.

Going forward, the aim should be to make the necessary adjustments to account for the range of capabilities at the state and local level, the relationships, and the complexity of the service delivery system and enact new policies and practices.

Natalie Helbig, Senior Program Associate Anthony Cresswell, Senior Fellow

www.ctg.albany.edu