CENTER FOR HUMAN SERVICES RESEARCH
UNIVERSITY AT ALBANY State University of New York

Integration of Early Childhood Data in New York State:

Examples of Successful Data Integration Efforts at the Local Level

Conducted by

The Center for Human Services Research

December 7, 2021

Introduction

Background

An Early Childhood Integrated Data System (ECIDS) collects, integrates, maintains, stores, and reports on information from the many programs comprising the early childhood care and education system within a state. A fully functioning ECIDS includes data about individual children and their families; about early childhood programs, providers, and services; and about the early childhood workforce. ECIDS data can thus be used to address program and policy questions that cannot be otherwise answered using a single data source; for example, an ECIDS would give policymakers a more complete picture of who receives specific services, and whether or not these services promote positive short- and long-term outcomes.

Case Example of Importance of ECIDS: Early care data considerations during COVID-19

The onset of the COVID-19 pandemic, and its impact on young children and families put into stark relief the need for data that illucidate service needs, both during the pandemic and beyond.

According to a 2020 report released by Child Trends, data is a necessity in order for both state and local early childhood leaders to address two basic needs: 1) ensuring that services are available where families need them, and 2) ensuring that families are able to access those services.[1] Additionally, Bui Lin and Richards (2020) underscored the heightened necessity of reliable and continuous data during the crisis to understand where services have been interrupted, where services are most needed, and how to best deploy resources.[2]

In New York State, there has historically been a lack of data and a limited ability to integrate data from multiple state and local agencies. These limitations left New York State with little ability to understand, at a state level, how families with young children were impacted by the closing of programs caring and providing supportive services for children ages birth-5 during the COVID-19 pandemic.

To address the immediate need to understand the need and availability of childcare during the pandemic, the New York State Office for Children and Family Services posted two weekly surveys (one for parents and one for providers) on their website beginning in April 2020. Parents were able to complete the survey indicating their childcare needs and programs were able to indicate any spots they had available. The data from this survey provided programs with information about the need for care, and families the availability of care in their area. This

---

[1] Data Quality Campaign (2020). Responding to COVID-19 and building for the future. https://files.eric.ed.gov/fulltext/ED607228.pdf

[2] Bui Lin, V-K. and Richards, D.E. (June 2, 2020). Integrated data can help states better respond to and recover from crises like COVID-19. https://www.childtrends.org/blog/integrated-data-can-help-states-better-respond-to-and-recover-from-crises-like-covid-19

effort was limited in scope and time, but it did successfully bridge a very limited data gap in the short term.

On a larger scope and scale, however, a statewide ECIDS could have enabled to be readily accessible to child-serving agencies across the state during a time when quick and accurate data was very much needed. An ECIDS would have allowed integration and analysis of ongoing data related not only to the need and availability of childcare but to the multitude of other service needs such as early intervention, home visiting, medical care and so on, and provided more comprehensive reporting from the start of the pandemic.

The Present Work

New York State would strongly benefit from an ECIDS, both to track the ongoing impact of the COVID-19 pandemic and to provide insight beyond the pandemic; currently, no such system is in place. In 2020, the Center for Human Services Research (CHSR) conducted a feasibility study on the development of a statewide ECIDS in New York State. The study found that while an ECIDS is feasible, development of one would face several challenges, including data sharing amidst privacy and confidentiality concerns, matching children across systems without a single unique identifier, and finding funding to implement such a comprehensive effort. CHSR proposed a pilot study to match children across two early childhood care and education systems; unfortunately, this project has been put on hold due to the pandemic.

Meanwhile, to continue building on the existing knowledge base regarding a New York State ECIDS, CHSR examined successful efforts to integrate data at a more local level to better understand the strengths and challenges of each system in order to develop recommendations and lessons learned to assist statewide discussions. This report summarizes data collected from interviews conducted with key informants across New York State who are currently using integrated data on a local or regional level to inform the planning and development of services to support the state's youngest residents.

Differences Between State-Level Integrated Systems and Localized Systems

There are several key differences in the organization, execution, and goals of a statewide versus local ECIDS that must be kept in mind when reviewing local system input. They include:

- obtaining informed consent from individuals is much easier at the local level since direct contact with families is a regular occurrence;
- locally integrated systems are better able to match children across systems, either with a unique ID or using other identifiers;
- locally integrated systems tend to be focused primarily on individual child outcomes and care management, not a larger picture or trends across time; and

- financial resources are typically more limited at the local level, which can hinder costly data system customizations as well as connections through Application Programming Interfaces (APIs) and similar data linkages.

These differences may make generalization of the lessons learned at the local level to a statewide system difficult. However, information about the utility of integrated systems, their potential benefits, and the implementation challenges can still inform the potential development of a statewide ECIDS.

Methods

The New York State Council on Children and Families (CCF) identified and recruited for interviews with three regional programs that used data integration programs and software to gather and analyze child and program data; an additional program was recruited by CHSR. For the purposes of these interviews, "data integration" was loosely defined as any system that brings data in from more than one source and can run basic reports. The programs are located in Western and Central New York and on Long Island. Programs in Western New York and Long Island are city- or county-wide databases; the Central New York system includes a wider multi-county geographic area. CHSR conducted four Zoom interviews with six individuals who were representatives from four different programs.

A semi-structured interview protocol was developed (see Appendix A). Questions focused on successes and challenges during different phases of system development:

- early discussions and conceptualization,
- planning and development,
- system implementation, and
- system maintenance and updating.

Telephone interviews occurred during January 2021. Researchers categorized summary notes from all interviews into a data matrix for the purpose of identifying themes.

Descriptions of the Data Systems in Use

All the systems interviewed are using one of two commercially available software programs to collect and store data.

The first, System for Tracking Access and Referrals (STAR), is utilized by many Help Me Grow programs across the country. This system is an online data collection portal that call center staff use to collect information about families, developmental issues and concerns, referrals to services, care coordination, developmental screening, outreach, and, to a lesser extent,

outcomes.[3] Outcomes typically focus on whether referrals were followedup on by the parent/family member and/or provider/program. Key informants reported that the end users they routinely speak to like this system because it is simple to use, flexible, and has prompt and helpful technical assistance. However, it typically only includes data input by one source. Data integration occurs when Ages and Stages Questionnaire (ASQ) screenings are entered into a separate system which links with STAR; staff using STAR can view ASQ data entered by those outside Help Me Grow.

The second system is the COMET system, utilized by the Children's Institute in Rochester, New York and its technology partner SophiTEC. The COMET system is a customized data portal that collects individual- and program-level data and allows real-time reporting on outcomes. Unlike the STAR system, external programs can input assessment and screening data directly into the system, enabling data sharing between programs regarding a particular child.

Results
Interview responses were coded, analyzed, and organized by theme. Results are presented below.

Benefits to Integrating Data at the Local Level
All key informants described a number of benefits to integrating data at the local level. Some of these benefits can be applied at the state level, while others are relegated to a local system only.

Family consent ensured. The system includes some form of a built-in consent form, which ensures that included families agree to have their data shared. Families previously completed a paper consent form for all programs, but when COVID-19 forced programs to provide services remotely, programs had to switch to a verbal consent form. None of the respondents reported any issues with moving to a verbal consent system. This consent process is not particularly relevant to a statewide system due to a number of privacy and other legal concerns with implementing such a system, in addition to the time and effort it would take to build in such a process at the state level.

Faster report production. Each data system provides the ability to set up and run "canned" reports that pull in, merge, and export data from multiple sources. Additionally, the software has the flexibility to program specialized reports that can be saved and run as frequently as desired. The ability to pull data from multiple sources and automate output allows for much faster reporting than with a paper or spreadsheet system. It also allows programs to use

---

[3] Hernandez, R., Ginter, L., and Sutka, J. (2017). Capturing Help Me Grow's Impact with a Simple Click: STAR (System for Tracking Access and Referrals). https://helpmegrownational.org/wp-content/uploads/2017/01/STARWebinar_OrangeCountyHMG_12.11.17.pdf

current data in real-time quality improvement efforts: programs can print reports and use data with both internal and external program staff to monitor and improve service delivery. Faster report production is a key benefit to integrating data at the state level. Preparing reports at the state level often requires manual compilation of data from disparate state agencies, which takes time and can lead to data errors. Having these data automatically compiled into reports saves both time and the need for re-entry and analysis of data by hand.

More complete picture of child's strengths and needs. One program emphasized the importance of having all child-level data in one place in order to comprehensively understand a child's situation. This program stated, "*Data is used constantly by partners in the field, especially when doing screening. It's very important to get data into the hands of professionals who are working with the child. Integrated data systems help us do that.*"

While consistent data entry can pose a challenge, all programs that use these systems are able to share child-level data, including developmental screenings, medical diagnoses, and other important indicators, with all local partners and identify community-wide patterns.

Use of data for quality improvement purposes. As stated above, continuous, shared data collection allows programs and communities to identify where services are being delivered appropriately and where there are needs and gaps. One respondent specifically emphasized the value of having data available when meeting with external programs to follow up on referrals and service delivery, stating that "*The thrust of our work is using data for continuous quality improvement. The goal is not just to point out where there are needs, but to take action steps to improve the system and make a difference.*"

Allows programs to identify the populations they are reaching (and those they aren't). Pulling data from multiple sources allows service providers to better understand what supports are reaching certain children and families, and what services are not. For example, one key informant reported that when they look at program referral data (both referrals received and made from multiple agencies connected by the data system), they can analyze the data to see what groups of families are being connected to services, and what referrals are a dead end. Dead-end referrals trigger a quality improvement process that attempts to understand and address the cause, and eventually lead to a better service connection process. Understanding population needs in real-time has become crucial during the pandemic. For example, a key informant reported that at the beginning of the COVID-19 shutdown it was important for programs to be able to quickly identify groups of children who were not participating in remote or virtual services. Due to their data systems being connected, they were able to pull both individual-level and aggregate data from their systems to identify where they needed to provide either additional technological or other supports to ensure children continued to receive the services they needed. Beyond COVID, access to integrated data allows program

leaders to make decisions about outreach and referral sources quickly, which in turn allows programs to be more responsive to the changing needs of their target populations.

Customizability. Both software programs are flexible. While the cost or degree of customization varies, all respondents were satisfied with what data the software could capture and the reporting features offered. This aspect is important; because programs are constantly shifting and building on the services they provide, any software used must have the ability to add and modify data and fields as needed.

Data Integration Challenges

Key informants all described a number of challenges in either setting up, using, or updating the system. These include:

- Data definitions are inconsistent across regions, which makes integrating data beyond the local level very difficult. For example, even though four Help Me Grow programs in New York State use the same data management software, they define key terms and phrases differently. Terms such as "being connected" mean different things to different Help Me Grow sites. One site may define "being connected" as having been referred to a program or service, while another site may define the same term as the family actively participating in a program or service.   These different definitions mean identification of larger trends and implications cannot be accurately tracked.
- The quality of data going into the system is often questionable. Much of this data is self-reported from other programs, and there have been challenges in getting programs to provide consistent data.
- Sufficient data are not integrated into the systems to tell a full picture of the larger early childhood system in an area. The amount of data being brought in from external program sources is minimal. For example, Early Intervention and home visiting data are not integrated, so the usefulness of collected data is limited to individual- and program-level decision-making.
- No systems included in this report are currently tracking or linking early care and education workforce data comprehensively. While systems such as the New York State Education Department's TEACH system for early childhood educators and New York Works for Children's ASPIRE Registry do collect data on workforce development, these data are not currently linked to either the STAR or COMET systems or to one another. Workforce data is key in hiring, retention, and professional development tracking efforts. Linking this type of data is important to fully understanding the early childhood care and education system.

All key informants were satisfied with the current functioning and capabilities of their data system, but lamented the lack of integrated data at the state level. All respondents recognized the need to not only understand the populations they serve, but how those populations relate to the larger early care and education and support service systems. Respondents suggested that as New York State moves forward, there are a few key considerations they must keep in mind:

- The system has to be built with the end user in mind. Systems that are overly complicated or cumbersome will not be used extensively and will therefore be limited in scope and usefulness.
- Key policy questions should be identified so the system is built to answer those questions and thus meet the needs of state leaders. As one key informant observed, "*Before developing or choosing a system, it's important to identify key policy questions you want the data to answer. Otherwise, the data serves no purpose.*"
- A data committee or workgroup needs to include the "right" people. All respondents identified a key group of individuals that led the development and implementation components of their data system. These groups were comprised of both decision-makers and those with technical knowledge of both the software systems and the local programs. One key informant noted, "*We had an organization leadership team that endorsed and pushed for a data system,*" which allowed them to be successful in implementing one.

Summary and Next Steps

While the value of integrated data at the early childhood level has been recognized for years, COVID-19 has underscored that need. Some New York State agencies simply lacked the data they needed to understand the impact COVID-19 had on the children and families they serve, and thus needed to scramble to create new data tracking methods. If a statewide ECIDS had existed in New York during the pandemic, agencies could have better and more quickly understood the needs of young children and their families during this time. Moving forward, New York should use the lessons learned at the local level to guide development of a statewide data system that can provide the information needed to make informed and quick decisions when the situation demands it.

To that end, the New York State Early Childhood Advisory Council has created a workgroup to continue to work toward recommending a statewide ECIDS. Early meetings have focused on gathering information on the major data systems already collecting, storing, and analyzing different data within the New York State early childhood care and education system. Workgroup members have also met with national ECIDS experts and will meet with representatives from states that have successfully created integrated data systems. Examining

the successes and challenges of integrating data at a local level will help to inform this workgroup as it gains a comprehensive understanding of different local and statewide data systems. With help and support from the New York State Council on Children and Families and the Center for Human Services Research, the workgroup will continue to meet to further identify the need, value, barriers, and next steps to develop and implement an integrated early childhood data system to better support the state's young children and families.

Appendix.

ECIDS Key Informant Interview Questions

PRE-PLANNING/EXPLORING
- Can you describe your data collection, storage, and sharing methods <u>before</u> you developed an integrated data system?
- How did the idea for an integrated data system begin at your organization?
    - What barriers needed to be addressed at this point in the process? How did you address them?
    - What went particularly well during this phase?
    - What was the key lesson (or lessons) that you learned during this phase?
- What were the key research and/or policy questions the data system needed to answer?
- What were the logistical issues the data system needed to address?
- Who were the key individuals involved in the discussion and what role did they play?
    - Did a group of people form a data committee to lead the process or was it one or two key individuals?
- How did cost factor in to initial discussions?

PLANNING
- What did the planning process for the data system entail?
    - How long did it take before development began?
    - What barriers needed to be addressed at this point in the process? How did you address them? (e.g., cost, lead agency, etc.)
    - What went particularly well during this phase?
    - What was the key lesson (or lessons) that you learned during this phase?
- How did the planning process for the data system address the following:
    - What model to use (federated versus centralized model)
    - Types of data to integrate (e.g., child-level, program level, workforce level)
    - Consent to share identifiable data
    - Confidentiality/security (e.g., unique identifier)
    - Compliance (with FERPA, IDEA, HIPAA, Head Start Performance Standards, state/local laws)
    - Data Sharing Agreements
    - Matching Data
    - Identifying partners (management and technical partners within each agency (data owners))
    - Data governance (establishing structure (individuals/roles/responsibilities/accountability, decision-making processes, protocols, manuals, training, etc.)

DEVELOPMENT
- Can you briefly describe the development process of the data system?
    - If you decided to purchase a system that was commercially available, why?
        - How did you decide on the particular system? (e.g., cost, ease of use, practicality, etc.)
    - How long did development take?
    - What barriers needed to be addressed at this point in the process? How did you address them?

- o What went particularly well during this phase?
- o What was the key lesson (or lessons) that you learned during this phase?

IMPLEMENTATION

- When did you go live? How long did it take to develop and implement?
- What are the major uses of the system at this time?
  - o Examples:
    - Sharing information with stakeholders or policy makers
    - Developing standard reports
    - Answering key policy questions
    - Responding to external data requests
    - Evaluating early childhood program outcomes
    - Accountability and compliance
    - Sharing information with parents
- Who are the system users? Do they have different permissions within the system?

MAINTAINING/UPDATING

- Now that the system is functioning, can you describe how it is maintained and updated?
  - o Is there an ongoing process to address system needs?
  - o Who is responsible for maintaining the system?
  - o Who provides training and support for users?
  - o What are they ongoing barriers that need to be addressed?
  - o What has gone, or continues to go, particularly well?
- How has the system answered the key questions that you initially developed?
- How has its use changed over time?
- What are the overall lessons you have learned from this process?