CCRY Research Initiative Report

Background

The Communities Collaborating to Reconnect Youth (CCRY) Network is a group of workforce and youth development professionals, working in communities across the country.[1] One of the questions that the CCRY Network is trying to answer is; can the research efforts by various agencies within the Network be combined to provide some degree of evidence on what type of programming is effective in producing positive outcomes for out of school youth. In an attempt to answer this question, the Network started a research initiative with agencies in three metropolitan areas in the Network: Baltimore (*Mayor's Office of Employment Development*), Hartford (*Our Piece of the Pie* and *Capital Workforce Partners*), and Philadelphia (*Philadelphia Youth Network*). These areas and agencies were chosen because they all have some research and evaluation capacity, thus have datasets that can be used in an analysis to answer questions about demonstrated effectiveness of youth programming.

The research initiative started on April 23, 2012 with a conference call of representatives from the various agencies along with additional CCRY leadership. The main purpose of the initial conference call was to clarify the goals of the initiative as well as determine the next steps forward. The end deliverable of the research initiative would be a report that would outline the steps of the research initiative along with any findings that may be obtained from the combined datasets of the various participating agencies.

In proceeding forward, the research initiative undertook the following broad steps:

A. Creating a shared understanding of the goals of the report to be produced by the initiative
B. Identifying what are the common data elements across agencies that would fulfill these goals
C. Obtaining the data elements identified
D. Data translation and combining into a unified dataset for analysis
E. Data analysis of combined data set to answer the goals of the report
F. Producing the research initiative report for review and feedback by participating agencies and CCRY leadership
G. Report dissemination

A. Creating a shared understanding of the goals of the report to be produced by the initiative

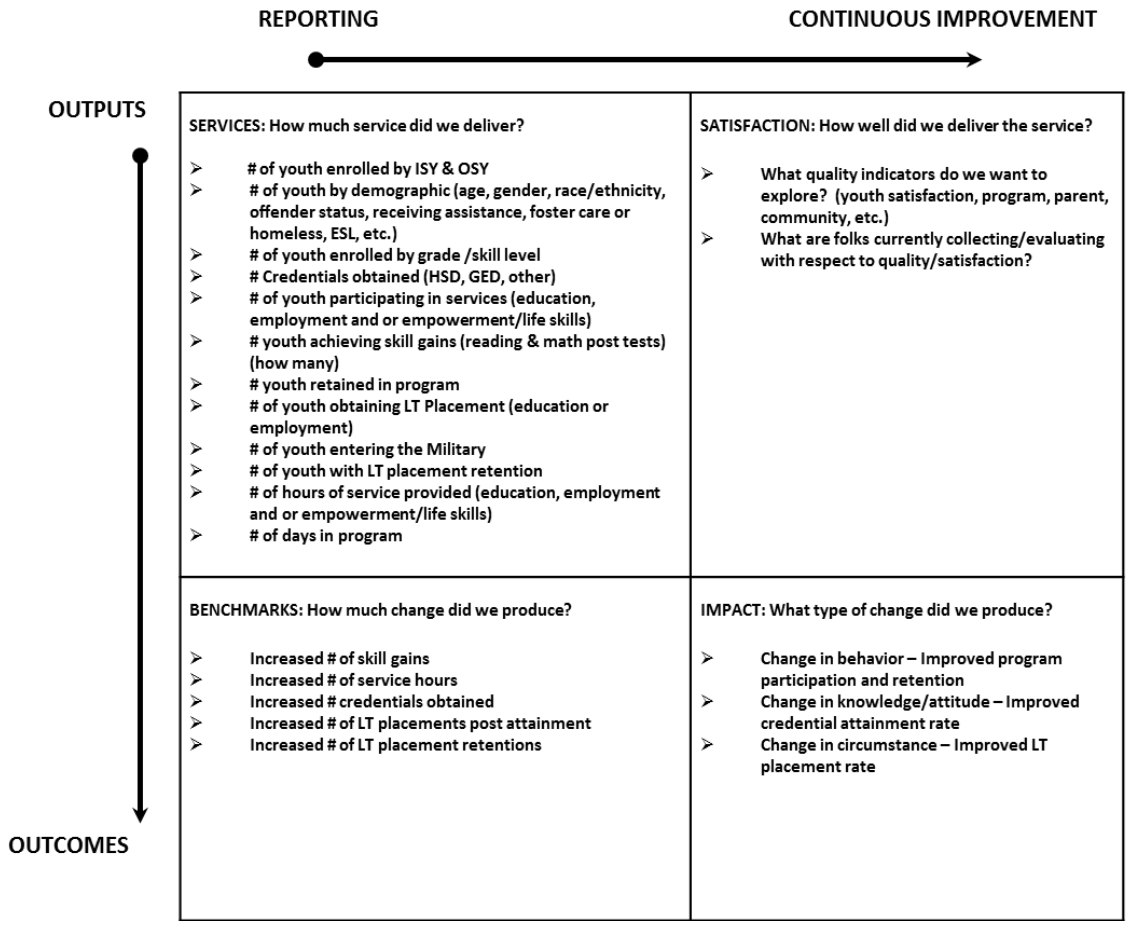
The research initiative group met several times to discuss what should be the goals of the initiative and the final report. Three goals were identified:

[1] http://www.ccrynetwork.org/about

1. To outline the process of data extraction and compilation across various agencies to answer assorted research questions, such as demonstrated effectiveness.
2. To provide a profile of the youth served, programmatic efforts, and outcomes achieved by various agencies.
3. To determine and document any demonstrated effectiveness that can be obtained by looking at the combined datasets of participating agencies.

B. Identifying what are the common data elements across agencies that would fulfill these goals

The next step of the initiative was to determine what data elements to be obtained from each agency. The group initially started with an ideal list of data elements to be obtained from each agency. For example, with regards to Goal 2 (an agency profile), the group identified questions such as: average number of service hours required to gain a skill, and average age of youth served. One of the tools we used to help us think about what the ideal list of data elements should be was Results Based Accountability (RBA). RBA allows for understanding of the types of data needed to accomplish the goals of an initiative. In using RBA, we considered questions around the RBA topics of Effort and Effect. Effort asks questions about how much effort was expended and how well was that effort delivered. Effect asks questions about how much change was accomplished and how well was the type of change produced. Below is a diagram showing the types of questions generated using this method.



The list of ideal data elements got to be rather lengthy and difficult to manage. We re-strategized and organized the data elements in bigger buckets or categories. The four big categories were Demographics, Program Services/Dosage, Short-Term Outcomes, and Intermediate/Long-Term Outcomes. Participating agencies were asked to identify what data elements they have in each of the four categories. Once the feedback was obtained from all participating agencies, each list of data elements were reviewed to determine if there were consistent elements across each of the lists. A final list was constructed based on these common elements including the format of the data elements. The list is below:

Name of Field/Variable	Variable Dimensions/Levels
Demographics	
Date of Birth	
Gender	
Race/Ethnicity	
Zip Code	
Parent	Yes/No
Foster Care	Yes/No
Homeless/Living in Shelter	Yes/No
Disability (Physical/Mental)	Yes/No
Juvenile Justice/Court Involved	Yes/No
OSY w/HSD	Yes/No
OSY w/GED	Yes/No
OSY w/o HSD or GED	Yes/No
Program Services/Dosage	
Education	
Reading Remediation	Time Spent in Minutes
Math Remediation	Time Spent in Minutes
GED Prep/Secondary/Alt Secondary Supports	Time Spent in Minutes
Post Secondary Transition Support/College Prep	Time Spent in Minutes
Employment/Workforce	
Job Readiness Training	Time Spent in Minutes
Internship/Work Experience (paid/unpaid)	Time Spent in Minutes
Occupational Skills Training (Vocational Training)	Time Spent in Minutes
Empowerment/Youth Development	
Civic Engagement/Community/Volunteer	Time Spent in Minutes
Service Learning	Time Spent in Minutes

Leadership Development	Time Spent in Minutes
Financial Literacy	Time Spent in Minutes
Conflict Resolution	Time Spent in Minutes
Mentoring	Time Spent in Minutes
Case Management/Follow-Up/Barriers	
Child or Dependent Care	Yes/No
Medical/Insurance	Yes/No
Transportation	Yes/No
Housing	Yes/No
Job Assistance	Yes/No
Peer Support	Yes/No
Outcomes	
Attained HSD	Yes/No
Attained GED	Yes/No
Attained Occupational Skills/Advanced Training Credential/Vocational Certificate	Yes/No
Attained Post-Secondary Credential	Yes/No
Literacy Gains (ISY and OSY)	Yes/No
Numeracy Gains (ISY and OSY)	Yes/No
Long Term Educational Placement/Enrolled in Post Secondary Setting (college, uni)	Yes/No
Long Term Employment Placement (been employed for at least 90 days)	Yes/No
Long Term Occupational Skills/Advanced Training Placement/Vocational Placement	Yes/No
Military Placement	Yes/No
Complete Educational level/class OR Grade to Grade Promotion	Yes/No
Completed JRT	Yes/No
Complete short-term subsidized internship	Yes/No

At this point, while we have successfully identified a list, one of the most pressing concerns is the combination of data from different agencies that do similar, yet slightly different types of programming. Therefore, when you consider things such as time spent on Job Readiness Training, what each agency does in terms of job readiness training may be vastly different from one another. While it is being combined into a single dataset to answer questions of effectiveness, the question about what exactly about job readiness training produces positive outcomes would require further research given the different programming involved across agencies. This is true of any of the other data elements included in the data set under program services.

C. Obtaining the data elements identified

Each participating agency was responsible for providing the data elements identified in Section B. Agencies provided anonymous, individualized student data. Not all agencies were able to provide all of the 45 data elements identified in Section B. The datasets were relatively large: Baltimore provided approximately 1,100 cases, Philadelphia provided approximately 1,800 cases, and Hartford provided approximately 400 cases. Each of the datasets was provided through Excel spreadsheets.

D. Data translation, and combining into a unified dataset for analysis

This section required that all of the datasets be combined in a single dataset for analysis. It involved understanding the data elements within each dataset, how it related to the master list of common data elements outlined in Section B, and translating the data to fit a unified model. The first step in the process involved creating a master key list which contained the elements identified in Section B, along with related data elements from each agency. An example is provided below:

Data Elements	Balitmore	Philadelphia	Hartford
Date of Birth	birth_date	Date of Birth	Date of birth
Gender	gender	Gender	GENDER
Race/Ethnicity	race_name	Race	Race
Zip Code	zip	Zip	zip
Living in Foster Care	Foster Care	Served Demos - Have you been or are you currently in living in a foster home	DCF Status
Homeless/Living in Shelter	Homeless	Are you homeless or living in a shelter	Barriers
Disability (Physical/Mental)	Column I - Disability Status	Do you have a disability medical conditions	Learning Disability & Physical Disability

A tentative list was constructed and distributed to participating agencies to ensure that the master key list correctly mapped the data elements in each data set to the list in Section B. Once the list was verified, the next task was to put each of the individual datasets into a single dataset. The single dataset is being constructed in the statistical software program, SPSS. Within each dataset, there was some amount of data translation that needed to be done. For example in Baltimore, several of the data elements provided often linked to a single data element from the list in Section B. For example, the element *GED Prep/Secondary/Alt Secondary Supports* for the final dataset is a combination of two elements in the Baltimore database (*YO(BLT)-GED Prep (YD/PP)* and *YO(BLT)-BCCC GED Prep (YD/PP)*). Or in another example, Hartford, the data is provided in multiple datasets, where each dataset can have multiple efforts of the same participant. This is a product of the way that the data management system exports its data.

Once data has been translated within each dataset, the data is then transferred over to the master file. In cases where there is categorical data, such as with the ethnicity/race data element, each categorical data element within each dataset needs to be translated into a unified set of categories. So for example, Baltimore had the following categories for ethnicity: *American Indian/Alaskan Native, Asian, Black/African American, Other, White*; Philadelphia had the categories: *African-American, Asian, Bi-Racial, Caucasian, Hispanic, Multi-Racial, and Other*; and Hartford had the categories: *African-American, Asian, Hispanic, Bi-Racial, and Multi-Racial*. In transferring the data, there needs to be a consistent translation of these categories across datasets.

This task of data translation and unifying requires significant time and effort to complete. The group is currently at this step.

E. Data analysis of combined data set to answer the goals of the report

Data analysis is geared towards fulfilling the last two goals of the initiative, namely:

2. To provide a profile of the youth served, programmatic efforts, and outcomes achieved by various agencies.
3. To determine and document any demonstrated effectiveness that can be obtained by looking at the combined datasets of participating agencies.

Data analysis will involve a combination of descriptive statistics, cross-tabulations, correlations, t-test/ANOVAs, and possibly a regression analysis. Descriptives and cross-tabulations will be used mainly to answer Goal 2. The other statistical procedures will be used to investigate any programming effectiveness common to all three agencies.

F. Producing the research initiative report for review and feedback by participating agencies and CCRY leadership

This report will be expanded to include the steps involved from Step E onward. An initial draft will be produced and then vetted by participating agencies and CCRY leadership to ensure accuracy and that it fulfills the objectives and expectations set forth by CCRY.

G. Report dissemination

A final copy of the report will be disseminated to the Network and others potentially interested stakeholders.