

**Evaluation and Research Support Services for the Pennsylvania Commission
on Crime and Delinquency (PCCD)**

**Report 2: Final Results of Performance Measures Audit for
Pennsylvania Commission on Crime and Delinquency**

March 2011

**Mary E. Poulin
Matthew R. Perkins
Jason Trask**



**Justice Research and Statistics Association
777 North Capitol St., NE
Suite 801
Washington, DC 20002**

Contents

Project Overview 2
 Project Phase 1 2
 Project Phase 2 4
Phase 2 Findings 5
 Revision of Erroneous or Poorly Worded Measures 5
 Performance Measure Categories 13
 Proposed Routine Performance Measure/Data Audit Strategies 19
 Using Performance Measure Data 21
 Ability to Answer Questions Posed by the Governor and Others 23
Conclusion 28

Project Overview

The Justice Research and Statistics Association (JRSA) received a two-year award to provide evaluation and research assistance to the Pennsylvania Commission on Crime and Delinquency (PCCD) in order to help PCCD improve its ability to produce meaningful information on the performance of PCCD-funded projects. The primary component under this collaborative evaluation/research support project was to develop a process for the selection, submission, analysis, and use of performance measurement data. The project also included guidance to PCCD on other research and evaluation support efforts.

The primary tasks of the project occurred in two phases. Phase One activities informed the activities of Phase Two. This report provides a summary of what occurred in phase one and a detailed discussion of the activities in Phase Two.

Project Phase 1

In the first phase of the project, we identified and assessed needs associated with the revision of PCCD's performance measurement approach. Our findings are available in the report entitled, "Report 1: Assessment of PCCD's Office of Criminal Justice System Improvements (OCJSI) Current Performance Measurement Approach" published in February 2009. Here is a summary of the Phase One findings:

- With Egrants, OCJSI's current performance measurement system has a strong framework overall to produce information on program performance, but a number of present practices diminish its power to produce a great volume of accurate, worthwhile information. Enhancements to the system from preparation/organization to analysis and reporting will help strengthen it. Of the needs and issues discussed, some can likely be

implemented with little cost or time and others will require significant effort and resources.

- Through Egrants, PCCD's online performance measurement and grantee monitoring system, OCJSI has tremendous infrastructure in place to collect and report data; however, infrastructure development is needed in other areas. According to our review of Egrants as well as discussions with PCCD staff, there are many errors in the data submitted in Egrants. The method for selecting performance measures, both required and optional, needs revision.
- As far as we know, PCCD does not have a strategic plan. A strategic plan would help PCCD justify the programs they seek to fund, aid in the selection of PCCD-required performance measures that fit with the goals and objectives of the strategic plan, and permit PCCD to assess how the program accomplishments reported in the Fact Sheets are helping PCCD meet the goals and objectives stated in the strategic plan.
- PCCD's performance measurement system has the capacity to answer a number of important questions about programs, but fine-tuning of the analyses and some additional data collection will go a long way in enhancing this.
- The reports that PCCD has issued to date have produced meaningful information and are easily accessible, but they can be improved to provide answers to a variety of questions. PCCD should add recommendations to help interpret report findings and put mechanisms in place to see if recommendations are implemented. Reports produced are easy to read and were produced in a timely fashion; however there are many funding streams without any reports. We are not aware of any external, independent review of the data or results.

Project Phase 2

In Phase Two of the project we completed an audit of Pennsylvania Commission on Crime and Delinquency (PCCD)-developed performance measures. This report describes the process and results of the audit. The audit included only those measures active as of September 2010 that were created by PCCD (n = 2,117) for the purpose of assessing the performance of grantees; it did not include measures developed by grantees. The objectives of the audit were to:

- Improve the data collection process,
- Aid in data verification, and
- Facilitate use of performance measures by PCCD and grantees.

A preliminary audit, which was completed in January 2010, included a review of active PCCD-defined performance measures, which involved grouping apparently similar performance measures into categories and naming those categories, identifying apparently duplicate performance measures, and identifying erroneous or poorly worded measures. At the conclusion of this preliminary audit, PCCD decided that JRSA should proceed with a more thorough audit that would include revising all the erroneous or poorly worded measures, adding grant-specific information to the measure (e.g., target, special definitions of key terms, in some instances time frame), placing measures in multiple categories (as needed), identifying routine data audit strategies, proposing ways to facilitate the use of performance measure data by PCCD and grantees, and assessing the extent to which the current performance measures permit PCCD to respond well to frequently asked questions received from the Governor regarding program performance. JRSA carried out these tasks requested as a result of the preliminary audit. It is important to note what JRSA did not do for the audit; the measures were not examined as a whole for each program/funding stream to assess the extent to which, together, the measures are

able to provide a good picture of program accomplishments. This is an important task, but was beyond the scope of the audit. In order to accomplish this, it is necessary to have an understanding of how a program/project is designed, either through a logic model or other tool that explains the program model.

Phase 2 Findings

This report provides several recommendations and proposes a number of strategies to help PCCD assess the performance of grantees and use the performance measure data collected. We make these recommendations in light of what we know is an ever-changing environment regarding the performance measures collected, why they are collected, and how they are used. Circumstances beyond PCCD's control may make it difficult to implement some of these recommendations (and it is possible that some are already in place), but we believe that implementation of these recommendations will significantly improve PCCD's performance measurement approach.

Revision of Erroneous or Poorly Worded Measures

We reviewed and proposed revisions, as necessary, to each of the 2,117 measures that PCCD provided to us in an Excel file. Revisions were made to the Excel file, but examples of proposed revisions are provided in this report. The measures were organized by PCCD offices: Courts, Corrections, Juvenile (Juvenile Accountability Block Grants [JABG]), Juvenile Evidence-Based, Juvenile), Law Enforcement, Victim Services, Technology, Weed and Seed, Training, Evaluation, and the Office of Criminal Justice System Improvement (Justice Assistance Grants (JAG), Generic). Measures were reviewed and revised as needed to address concerns regarding precision (e.g., reporting time frame), how to carry out calculation(s) required to report the measure, the need for definition of key terms, and to ensure that they meet

the specifications of the Egrants. The technical specifications of Egrants, PCCD's reporting system, demands that the response be presented in numeric form and only one response be entered per measure. When a revision was proposed, it was placed in the "new proposed language" heading with the proposed revision(s) in red text. In some instances, questions remained about how to revise the measure, or the measure required revision but JRSA lacked information to permit proposing a revision. When this occurred, a comment or question was placed in the "Notes" column. PCCD should review and address these prior to proceeding with use of the measure.

A good performance measure will have a great deal of specificity in its wording. The wording should indicate the subject of the measure; how the measure will be collected; when the measure will be collected; in what format the data will be collected; and, in cases when calculations are required, the formula to conduct the calculation. Though it may seem redundant, the clarity can prevent misunderstandings or mistakes that affect one's ability to understand program implementation or outcomes. The following is an example of a measure that contained some ambiguity:

*Of the offenders who completed the program, the number who remained drug-free during the aftercare program during the reporting period.
Documentation must be maintained for future monitoring purposes.*

A number of questions can be raised from the current wording. Is the aftercare program considered part of the overall program? Will individuals whose participation in the aftercare program crossed over from one reporting period to the next be counted in each reporting period? How will drug-free be measured? Self-report? Urinalysis? To bring greater specificity and to insure that participants are counted only once, a one-time marker (completion of the aftercare

program) was recommended as key for inclusion. Note that, even with revisions, this measure will not capture drug-free status for individuals who never completed the aftercare program.

*Of the offenders who completed the program **and, during the reporting period, completed the aftercare program**, the number who remained drug-free [how is drug-free measured?] during the aftercare program. Documentation must be maintained for future monitoring purposes.*

A critical step in insuring that a measure is well-defined is an examination of how the measure can be misinterpreted. It is a common mistake to create measures that have unwritten assumptions, usually because the meaning is clear to the writer or because the terminology will be clear to the intended groups. However, as time goes by, such shared knowledge can be lost and new parties introduced to the reporting process. Sometimes there is an insufficient examination of the possible interpretations of a measure. Here is an example of a performance measure with this problem:

YAVRP--# of combined visits with probation staff

Here is how we propose to revise the measure:

Number of combined visits (to whom?) between law enforcement and probation staff this reporting period.

The performance measure was revised to address concerns regarding with whom the probation staff will conduct visits and when it will be reported. It still does not define YAVRP, though we believe that this is the acronym for the program for which this measure is collected. Further, it remains unclear who is being visited. This must still be addressed.

Reviewing the measures can help in fixing mistakes or inconsistencies in the measure.

The following measure shows an example of a common mistake:

The percent of program youth who have exhibited a change in targeted mental health issues. Self-report or staff ratings are most likely data sources. A. Number of program youth with the noted behavioral mental health

improvement discharged this reporting period. B. Number of youth discharged this reporting period. C. =Percent (A/B).

In this measure a percentage is requested, but the formula expressed (A/B) is a ratio where the proper formula for expression as a percentage would be $((A/B)*100)$. The assumption by the writer may have been that in order to express the relationship of A to B as a percentage the ratio would be multiplied by 100, but proper expression can reduce the opportunity for error.

Beyond these examples of measures that require revision, we identified a number of themes regarding proposed revisions to the measures. Some of these themes were more common in the measures used by some offices than in others. The themes included:

- the data collection period,
- need for specificity in what is to be measured,
- replacing symbols and acronyms,
- lack of information regarding the data source/how to count,
- the measure is not expressed as a numeric value,
- need to use neutral language, and
- the addition of formulas for calculations.

Examples of proposed revisions (in **bold** text) are provided below.

Data Collection Period

The most common change needed is to indicate a time period for which the performance measure applies and/or how often it is to be reported. PCCD should review proposed revisions to ensure that proposed time periods are appropriate for each measure. Tables 1 and 2 provide examples of measures to which this issue applies.

Table 1. Addition of Data Collection Period: Example 1

Original Measure	Revised Measure
Total number of individuals attending SFP training. This number should include those that were seeking certification as facilitators, as well as others observing or participating in part of the training, but not completing the training or becoming certified. Count each individual only once. Documentation must be maintained for future monitoring purposes.	Total number of individuals attending Strengthening Family Program (SFP) training during the reporting period . This number should include those who were seeking certification as facilitators, as well as others observing or participating in part of the training, but not completing the training or becoming certified. Count each individual only once. Documentation must be maintained for future monitoring purposes.

The second example raises another issue regarding the specification of relevant time periods and ambiguity resulting from the wording of the measure. As the example in Table 2 shows, the placement of the term ‘in the reporting period’ leads to confusion about whether or not it means that the arrest occurred during the reporting period or whether termination of the program occurred during the reporting period. The rewording clarifies the issue and should insure that individuals terminated from the program (during the reporting period) for a new drug (non-DUI) arrest during a prior reporting period are not left out of the measure. Under the original wording, this could happen either because a lag between arrest and termination, even if very short, crossed two reporting periods, or because there was a delay in a new arrest being reported/recognized.

Table 2. Addition of Data Collection Period: Example 2

Original Measure	Revised Measure
Total number of D&A RIP offenders terminated from the program due to a new drug (non-DUI) arrest in the reporting period. Documentation must be maintained for future monitoring purposes.	Total number of Drug and Alcohol Restrictive Intermediate Punishment (D&A RIP) offenders terminated from the program during the reporting period due to a new drug (non-DUI) arrest. Documentation must be maintained for future monitoring purposes.

Need for Specificity in What Is to Be Measured

Several measures do not offer enough specificity regarding the type of information to collect. Assuming that others will know how to interpret terms used in the measure can lead to an inability to compare results across grantees or to aggregate data. For example, for the measure in Table 3, there is ambiguity about what constitutes a “new youth.” The term could be interpreted as youths enrolled during the reporting period or youths enrolled before the reporting period who had not been previously matched with a mentor.

Table 3. Specificity in What Is Measured Example

Original Measure	Revised Measure
Number of youth matched with a mentor during the quarter. Include in this count only the new youth matched during the quarter. Documentation must be maintained for future monitoring purposes.	Number of youth matched with a mentor during the quarter. Include in this count only the new youth matched during the quarter. A new youth is a youth not previously matched with a mentor since program enrollment. Documentation must be maintained for future monitoring purposes.

Replacing Symbols and Acronyms

Replacing symbols and acronyms with full words will reduce the possibility of misunderstanding. See Table 4 for an example of this issue.

Table 4. Use of Symbols and Acronyms Example

Original Measure	Revised Measure
The number of DMC-related policies or procedures created, amended or rescinded during the reporting period. Policies or procedures can be developed at the state or local levels. A policy is a plan or specific course of action that guides the general goals and directives of the program or agency. Include policies that are either relevant to the topic area of the program or policies that affect program operations. Documentation must be maintained for future monitoring purposes.	The number of Disproportionate Minority Contact (DMC)-related policies or procedures created, amended or rescinded during the reporting period. Policies or procedures can be developed at the state or local levels. A policy is a plan or specific course of action that guides the general goals and directives of the program or agency. Include policies that are either relevant to the topic area of the program or policies that affect program operations. Documentation must be maintained for future monitoring purposes.

Lack of Information Regarding the Data Source/ How to Count

Measures require instructions on how to count the data (e.g., who should be included when counting “successful” completions?) or what sources of data should be used (e.g., self-report or official records?). Lack of instruction impedes cross-program comparisons and aggregation of data. Table 5 provides an example of this issue and a proposed revision to the measure.

Table 5. Defining How to Count Example

Original Measure	Revised Measure
The time measured in minutes between an event and contact via fax for the reporting period. Documentation must be maintained for future monitoring purposes.	The average length of time, in minutes, from an event to contact via a successfully transmitted fax during the reporting period. For attempts to contact individuals with fax numbers on file. Documentation must be maintained for future monitoring purposes.

Measure Not Expressed as a Numeric Value

Many of the measures are open-ended and do not allow for responses to be quantified. Egrants requires that measures be entered as numbers. Table 6 offers an example of this issue.

Table 6. Use of Numeric Values Example

Original Measure	Revised Measure
Completion of a plan for the growth and expansion of the agency	Were plans for the growth and expansion of the agency completed during the reporting period? If completed, enter '1'. If not completed, enter '0.'

Need to Use Neutral Language

The performance measure should reflect only what information is being collected. Two common concerns were that: 1) the desired outcome was stated in the performance measure (e.g. a measure stating “Five percent decline in...” rather than “Percentage change in...”) and 2) the measure assumes causality (e.g., a measure stating “The percentage change in X as a result of Y”

rather than “The percentage change in X following implementation of Y”). Below, in Table 7, is an example in which the measure reflects both of these concerns.

Table 7. Use of Neutral Language Example

Original Measure	Revised Measure
A 3% increase in the use of federal grand juries to develop compelling cases attacking gang violence from the previous grants	Percent of cases attacking gang violence in which federal grand juries were used to develop cases. A = number of cases of gang violence in which grand juries were used to develop the case during the reporting period. B = number of cases in which any method was used to develop a case to attack gang violence. C = ((A/B)*100).

The Addition of Formulas for Calculations

In instances in which the measure is complicated or a computation must be used, the formula for the calculation should be expressed as part of the measure. This both prevents confusion or inconsistency in the calculation of the performance measure and provides a guide to the data that must be captured in order to produce the performance measure. See Table 8 for an example of this.

Table 8. Add Formulas for Calculations Examples

Original Measure	Revised Measure
The percent of all students at La Academia who have increased their attendance at the Charter School. This will be reported quarterly.	For the PCCD reporting period, the percent of all students at La Academia who have increased their attendance at the Charter School this school quarter compared to last school quarter. To calculate: A. Number of students at La Academia who have increased their attendance at the Charter School this school quarter compared to last school quarter. B. Number of students at La Academia this school quarter. C. Percent change (((A/B)-1)*100).

Performance Measure Categories

Performance measures were reviewed and grouped according to the type of information that the measure is intended to collect. The performance measure categories were developed in order to help PCCD staff quickly identify pre-existing performance measures that would fit a new award, enhance PCCD's ability to report aggregate performance measure data across grantees and programs, and prevent the generation of duplicate measures. Each measure was placed into either one or two categories, depending on the purpose of the measure. When a new measure is developed, PCCD should determine the purpose(s) of the measure and place it in one or more categories. Some of the categories and definitions here have been revised from those developed for the preliminary audit.

Categories

1. **Individual Assessment** – Refers to the use of a tool or decision-making process to determine how to respond to an individual. These assessments are often conducted at intake or the beginning of a program and may use various instruments to assess needs, risks, or other issues. Assessment differs from evaluation; evaluation is used to determine the effectiveness of services and programs.

Example: The number of youth assessed as needing substance abuse treatment during the reporting period.
2. **Behavior** – Used to capture the actions or changes on the part of individuals. Often captures changes in behavior utilizing some form of pre- and post-test or survey.

Example: The number of program youth who show a decrease in antisocial behavior during the reporting period as determined by pre- and post-test.

3. **Case Processing** – Designed to capture data relating to the handling of people, actions, or decisions. These measures are often used by agencies such as courts, probation, and police departments to track people, actions, or decisions as they proceed through the system. These measures often include information pertaining to case disposition and processing time. When admittance or enrollment in a program does not refer to actual participation in the program, it should be included here.

Example: Report the average number of calendar days from a case being officially referred to pretrial services to the case being closed by the pretrial program during the quarter.

4. **Change or Improvement** – For the purpose of capturing data regarding desired outcomes in the course of or following the implementation of the program. Changes that are behavioral or captured in other categories such as crime, cost-savings, surveys, or system improvement should NOT be included in this category.

Example: Number of youth discharged this reporting period who have improved family relationships.

5. **Compliance Monitoring** – Used to collect data to determine whether or not the grantee is adhering to rules, regulations, and/or requirements.

Example: The number of establishments that were found to be noncompliant during Age Compliance Checks during the quarter.

6. **Coordination** – Designed to capture data that can be used to determine the degree to which agencies, groups, or other entities are working together to meet program goals.

Example: The number of collaborative members who sign a Memorandum of Understanding (MOU) as a commitment to the planning process during the reporting period.

7. **Cost-Savings** – Used to capture data that can demonstrate a reduction in the monetary cost related to a program or process.

Example: The number of incarceration days saved by participating in the Mental Health Court (MHC) program. To calculate, collect the number of days persons would have been sentenced, if not for mental health court participation, and subtract any days served for infractions.

8. **Crime** – Used to capture occurrences of crime. In addition to **general crime**, this category can include more specific subcategories, including: **violent crime, property crime, juvenile crime, drug crime, gun crime, and gang crime**.

Example: The total number of violent crimes within targeted community during this reporting period. Violent crime is defined as Uniform Crime Report (UCR) Part I offenses known to police. The data source for this measure is official police records.

9. **Employment & Employment Training** – Includes two subgroups: **employment** and **employment training**. The **employment** subgroup captures data directly related to getting, keeping, losing, or having a job, while the **employment training** subgroup captures data related to training and other services designed to improve or otherwise change the employability of those served.

Example 1: The number of Drug and Alcohol Restrictive Intermediate Punishment (D&A RIP) offenders who gained program-approved employment while participating in the

program during the reporting period. Only count those who were unemployed at their time of acceptance into the program.

Example 2: The number of program participants who completed a job training course during this reporting period.

10. **Materials/Reports/Products** – Used to collect data related to production and dissemination of items such as reports, presentations, public awareness ads, brochures, newsletters, etc.

Example: The number of direct mailings distributed during the reporting period.

11. **PCCD Tracking** – Appear to be related to PCCD’s oversight of grants. They seem to be internal measures. We are not sure why these are included in the list of performance measures.

Example: The amount of funds in whole dollars that are allocated at the local level during the reporting period. Program records are the preferred data source.

12. **Personnel** – Used to capture data related to program staff or human resources. These measures do not include employment data such as number of jobs created.

Example: The total number of officer overtime hours paid out during the quarter for age compliance checks related to Enforcing Underage Drinking Laws (EUDL).

13. **Policy Development** – Capture data pertaining to the creation, alteration, or enactment of policies.

Example: The number of alcohol-related school policies that were changed during the quarter. This is to include any new policies that were adopted during the quarter.

14. **Program Completion** – Capture graduation, termination, or other discharge statuses from a program. These measures are often used in combination with a program

participation measure so that the number of participants completing a program can be compared to the total number of program participants.

Example: The number of youth who have successfully completed all the necessary requirements of the program(s) or activities. Only count those youth who have successfully completed the program(s) or activities during the quarter.

15. **Program Operation** – Used to capture general data pertaining to the implementation or functioning of programs. This includes the subcategories of **implementation fidelity**, **meetings**, and **direct service**. Implementation fidelity refers specifically to those measures used to collect data related to how well the program is carried out according to plans. Direct service measures are designed to capture data related to the provision of services. Direct service measures focus on the service provided, not the number of people served.

Example: The number of new sites using the classroom setting component of The Incredible Years (TIY) Child Training Program. Only enter the number of sites that began using the curriculum during the reporting period.

Example: The number of Life Skills Training (LST) lessons delivered with a fidelity score of at least 75%. Fidelity is to be assessed by an LST- trained building administrator or evaluator utilizing the LST Implementation Checklist. Count only those lessons delivered during this reporting period.

Example: The number of social and recreational activities conducted for youth, their families, and mentors during this reporting period.

16. **Program Participation** – Used to collect data on involvement in the program by individuals or groups as well as the number of available program slots and program capacity. These measures are not used to capture data related to the provision of services. Example: The number of youth who participated in community policing youth activities (CPYA) during the reporting period. (For example: PAL, youth police academy, cops and kids programs).
17. **Recidivism** – Capture the reoffending of individuals. May include measures of rearrest, substance abuse, reincarceration, self-reported delinquency, or other measures. Example: The number of program graduates rearrested within one year of graduation for any offense during the reporting period.
18. **Substance Abuse** – Capture data pertaining to the use and/or abuse of drugs (illegal or legal prescription), alcohol, tobacco, etc. Example: The number of program youth who receive at least one positive drug test result during the quarter.
19. **Surveys** – Refers to any method used to gather information from individuals. These may be called surveys, questionnaires, pre- and post-tests, etc. This category also includes counts of surveys received or completed. Contains two subcategories: **Knowledge & Awareness** and **Customer Satisfaction**. Knowledge & Awareness is used for to demonstrate change in learning or understanding. Customer Satisfaction is used to measure the degree to which participants were pleased with programs or services. Knowledge & Awareness Example: The number of individuals who reported increased confidence in their ability to recognize child abuse and/or neglect. These data are to be

collected using completed pre- and post- training surveys for individuals who completed training during this reporting period.

Customer Satisfaction Example: The number of attendees who indicate satisfaction with the usefulness of the training provided. These data are to be collected using completed post-training surveys for individuals who completed training during this reporting period.

20. System Improvement & Efficiency – Designed to capture the impact that changes have on the operation of programs, services, and/or systems. These changes are often intended to enhance the capabilities or functioning of these programs, services, and systems.

Example: Report the raw number of case files (not individual youth) that are completely automated (i.e., all required data about the case are entered in the automated system and ready for use) during the quarter.

21. Training & Technical Assistance –Designed to capture data pertaining to classes taught or attended, and/or practices delivered or received to enhance the ability of program personnel to perform a task or operate a program. If training is part of a service being provided to a program participant (e.g., number of youths attending training on effects of drug use), do NOT include it here.

Example 1: The total number of agency staff attending refresher training during the reporting period.

Example 2: Total number of technical assistance requests fulfilled during the given quarter.

Proposed Routine Performance Measure/Data Audit Strategies

The performance measures used are not static. Over time, measures will be modified, some will be deleted, and new ones will be added. Grantees and programs/projects funded are

dynamic as well. Consequently, it is important to ensure that new measures developed have well-worded descriptions and that the data submitted are accurate. There are two key points at which audit strategies should be employed: when measures are being identified for a solicitation or program and following the end of the first reporting period.

When performance measures are being identified for a solicitation or program, several steps should be taken. Before developing new measures, the person(s) responsible for identifying measures should search Egrants, by category, to determine whether preexisting measures are appropriate or can be modified to meet the solicitation or program needs. To ensure that measures are well-worded, we recommend that several sources be consulted:

- Person(s) developing the measures should check the measures to make sure that they: state time frame for data included, request only numeric data, include only one number per measure, define all terms, select appropriate category(ies), provide directions on how to do any required calculations, and spell out any acronyms included in the measure;
- PCCD staff with experience in performance measurement should review the measures to identify any concerns; and
- Previous and/or potential grantees working in the topic area of the solicitation should review the measures to identify concerns about the measures, such as potential challenges with collecting the data and how to define key terms used in the measures.

The end of the first reporting period is a critical time to audit data because it is the first opportunity to check actual data to determine how the measure is being interpreted by grantees. Each grantee's data should be reviewed closely at this time to identify possible problems. This process should include reviewing each grantee's proposal to assess whether proposed activities and objectives are in line with submitted data, and contacting a few randomly selected grantees

to ask them how they interpreted the measure and what process they used for collecting data for the measure. These two steps will help determine whether revision of a measure is needed.

Using Performance Measure Data

Performance measure data collected by PCCD can be used for three major purposes: to assess an individual grantee's performance, to gauge the overall performance of a particular funding stream or group of programs/projects, and to consider PCCD's contribution to criminal justice system operations and objectives in Pennsylvania. The performance measure data can be used to answer a wide variety of questions for these purposes, but there are limitations to their uses. These data can help answer questions regarding the implementation of programs/projects and whether the programs or projects accomplished their objectives, but generally they cannot be used to determine whether the program/project was responsible for (i.e., caused) a particular outcome. A program evaluation with a rigorous research design is needed to answer these types of questions. Further, the data cannot be used to track the performance of a particular individual served by a program/project because only program-level data are collected. Finally, it is not possible to conduct any analyses that require crosstabulations of two measures. For example, if there is one measure recording the number of persons successfully completing the program and another measure recording the number of persons who recidivated, it is not possible to report on the number of persons successfully completing the program who recidivated. It would only be possible to report on this later concept if there were a measure explicitly asking for the number of persons successfully completing the program who recidivated.

The performance of individual grantees should be assessed throughout the life of the award. Process measure data, which collect information on the implementation of the program, should be reviewed at the end of each reporting period, with a particular focus at the end of the

first year. At this stage, the purpose is to help the grantee, if needed, with carrying out planned program activities. PCCD should select a threshold for determining whether implementation is problematic, and provide feedback to grantees at the end of the first year to let them know how their program is doing. For example, PCCD might consider that program implementation is problematic when less than 75% of program activities are being carried out as planned. At the end of the award, PCCD should produce a standard report for each grantee that provides a summary of program accomplishments. Ideally, this report should include what the grantees hoped to accomplish (e.g., a list of program objectives) and what they actually accomplished (e.g., activities carried out and outcomes).

Reports developed for individual grantees at the end of their award can be useful to PCCD as well. Data from these reports can be aggregated to gauge the overall performance of particular funding streams or groups of programs/projects. PCCD can also examine each individual report to identify those programs/projects that were the most successful in implementing the program/project as planned and achieving desired outcomes for a particular funding stream. Information from individual and aggregate reports may then facilitate future funding decisions for individual grantees based on their performance and be used to develop guidelines for the design and implementation of subsequent programs/projects funded. Further, PCCD may decide not to fund certain types of programs based on these data or to require that funded programs/projects incorporate a particular approach to address a problem. Information from aggregated reports should be shared publicly in order to contribute to the knowledge base in the area. Finally, PCCD should use data from aggregated reports to decide whether a formal evaluation would be worthwhile.

Ability to Answer Questions Posed by the Governor and Others

As PCCD begins to use measures that cut across offices and funding streams, it will be in a better position to answer questions common to criminal and juvenile justice programs/projects. For example, when measures such as recidivism and the number of clients served are defined in the same way, PCCD will be able to aggregate these data to report on issues such as recidivism reduction occurring after the provision of PCCD-funded services.

An examination of the *2008-2009 Pennsylvania Governor's Report on State Performance* provides an opportunity to look at issues surrounding the use of PCCD performance measures and can help identify what measures should be added or revised to provide better information for the report. Performance measures are limited in what they can provide. This is of particular concern if the measures were developed for specific programs rather than for the "Key Objective" for which they are used in the Governor's report. Here again, we recognize that PCCD may face limitations in being able to address the issues identified here regarding the objectives and performance measures used in this report, but where possible, modifications will help PCCD provide better information for the Governor's report.

Among the Key Objectives listed in the Public Safety section of the report, six have data sources provided by PCCD. Here we offer our thoughts regarding how measures might be revised or what measures might be added in order to provide better data to assess whether the key objective has been accomplished.

Fingerprints

PCCD provides resources to enable the electronic capture and recording of fingerprints, hand impressions, and photographs for online submission to the Pennsylvania State Police. This is a difficult measure to analyze because the wording of the Key Objective is open to two

interpretations. One interpretation is that the “collection rate” of fingerprints and photographs refers to the rate of speed at which records are acquired by the State Police. The other possible interpretation relates to percentage of total [electronic] fingerprints and/or photographs captured by law enforcement in the state that are submitted to the State Police. Subsequent discussion in the report suggests that both interpretations are possible. In either case, the performance measures are not sufficient. Measures are needed that can capture the speed of submission as well as the number submitted compared to the number that should have been submitted.

Key Objective	Performance Measures
Increase the collection rate of arrestee fingerprints and photographs (p. 66)	Number of locations with electronic identification technology
	Percentage of fingerprints submitted electronically

Jail Costs

The Key Objective states that the intent is to “reduce county jail costs” while the report states that “alternatives to incarceration for nonviolent offenders save money by freeing up jail space for more serious or repeat offenders without building new prisons (p. 75).” In response to this objective, the performance measures used are data from two intermediate punishment programs, both of which provide percentages of offenders completing the program and the average jail days saved per offender.

A number of assumptions underlie the ability of these measures to inform progress toward the key objective. The measures need to more clearly specify how the number of days saved is calculated. Are participants in the intermediate punishment programs offenders that would otherwise be going to jail? Is jail space freed up by offenders in the intermediate punishment programs filled by other offenders – either offenders with more serious charges/convictions who would otherwise go to a more secure facility, or offenders with less

serious charges who would normally not go to jail? And are offenders who do not complete the program subject to serving their full sentence in jail when exiting the program? If so, calculations of jail days saved apply only to offenders completing the program. To ensure that the measure provides good information for the objective, these questions must be addressed.

Key Objective	Performance Measures
Reduce county jail costs by providing alternatives to incarceration for nonviolent offenders (p. 75)	Counties participating in intermediate punishment programs
	Average jail days saved per offender
	Percentage of offenders completing program
	Counties participating in Drug and Alcohol county Restrictive Intermediate Punishment programs
	Average jail days saved per offender
	Percentage of offenders completing program

Evidence-Based Juvenile Programs

In examining the use and efficacy of evidence-based programs for juvenile justice and delinquency prevention, the report contains six measures that relate directly to the objectives. Two of the measures relate to the use of research-based programs. Though the number of youth served may not capture how widespread the use of research-based programs was, it does show changes in the overall ‘use’ of such programs. One possible weakness to this measure is that although it provides the aggregate number of youth served, it does not provide information on how this relates to the number of eligible youth.

The four remaining measures address the relative effectiveness of the programs measured. By utilizing the percentages of youth and families with positive outcomes, the measures can provide useful year-to-year comparisons. An element that is missing is the ability to indicate whether these success rates are comparable to either youths in non-research-based

program or youths not involved in programs. One can determine if the results are better than in the past, but not if they are better than for other populations.

Key Objective	Performance Measures
Increase the use and efficacy of evidence-based juvenile justice and delinquency prevention programs (p. 75)	Number of youth served in intensive, research-based programs
	Number of youth served in non-intensive, research-based programs
	Percent of families served that functioned better
	Percent of youth with improved school attendance
	Percent of youth served receiving no new charges
	Percent of youth served avoiding residential treatment

Criminal Justice Advisory Boards

The Key Objective for Criminal Justice Advisory Boards (CJABs) to increase efficiency relies on a performance measure that provides little information. The number of counties participating only serves as a measure of CJAB use throughout the state and does not provide any information on whether efficiency was improved and, if so, how.

Key Objective	Performance Measures
Increase the number of county Criminal Justice Advisory Boards to improve justice system efficiency through interagency planning and collaboration (p. 76)	Counties Participating

Victim Compensation

The suitability of the measures used to determine timeliness and effectiveness of victim compensation efforts is mixed. The average time to process claims is a suitable measure to examine the time required for compensation and by using average time the measure is comparable regardless of the volume of requests. However, the effort required of victims to receive compensation is not measured as well here. One could maintain that the number of

claims submitted (though the measure is not stated as such) is a function of the effort required (i.e., if the effort is less, more claims are submitted), perhaps in interaction with the total amount paid (i.e., if claims were for a lower average value, lesser effort must make it ‘worth it’ to submit claims for lower amounts of money). A more direct measure of effort (e.g., time required to submit a claim) would be better, or even a measure that provided information on the relative likelihood of submitting a claim (e.g., proportion of eligible claimants submitting a claim).

Key Objective	Performance Measures
Reduce the time and effort required for victims to receive compensation (p. 82)	Claims submitted
	Amount Paid to or on Behalf of Victims
	Average Time to Process Claims

Victim Notifications

This final objective for which PCCD provides performance measures is supposed to help determine whether the amount of time to notify victims has been reduced and whether the process of notifying victims has been streamlined. The performance measures currently used should be revised in order to address this objective.

(Number of) Total new registrations does not reflect on the time for notifications and provides no information relating to the streamlining of notifications (though it may reflect any efforts to streamline registration processes). It is likely that the number of registrants is related primarily to the number of eligible victims and efforts to publicize the Statewide Automated Victims Information and Notification Service (SAVIN) system.

Numbers of notifications can be a measure of aggregate performance but does not provide the key information about whether the notifications sent were done more quickly or in a streamlined manner. An assumption might be made that notifications by e-mail are quicker than those by phone, but any such assumption would be balanced by the increased difficulty of

knowing when an e-mail notification is received, or if it is even delivered successfully. More appropriate measures might be the time to ‘successful notification’ by phone or the time to receipt of confirmation that a sent e-mail was viewed.

Key Objective	Performance Measures
Reduce the time and streamline the process of notifying victims of crime of the release, escape, or custody status change of an offender (p. 83)	Offender Status Notifications by Phone Offender Status Notification by E-mail Total New Registration (PA SAVIN)

Conclusion

With the current performance measurement system, PCCD is well-positioned to produce a wide variety of information on the performance of the programs and projects they fund. However, some elements related to the design of the system and the performance measures being used, make production of good quality information challenging. The performance measure audit described here was carried out to help PCCD address this.

The proposed revisions and recommendations provided in this report are intended to help PCCD meet these objectives: improve the data collection process, aid in data verification, and facilitate use of performance measures by PCCD and grantees. Whoever is tasked with managing the implementation of the proposed revisions and recommendations will have a significant amount of work to do. This will include, but is not limited to: reviewing and approving proposed changes to the performance measures, doing some redesign to Egrants, training staff and grantees on revisions, and setting up a mechanism for producing reports on program performance. This is a worthwhile investment because making changes will not only help meet the stated objectives of the audit, but it will make it easier to use the performance measurement system in Egrants for both PCCD staff and grantees, improve the quality of the data collected, and result in greater use of the performance measure data.