



# COMMUNITY SAFETY & JUSTICE ALLIANCES: STEP-BY-STEP

# LISC

## Collecting Data for a Safer Community

### A toolkit for community-based problem-solving teams: What information is needed and how to find it

This toolkit was developed to help local coalitions looking to improve community safety do a more robust job of collecting data—particularly Byrne Criminal Justice Innovation program (BCJI) teams. It provides ideas for alternative sources of information and lays out concrete steps for collection, so data is identified, collected and transferred in a timely fashion. It discusses challenges and pitfalls that you should be aware of and ends with an exercise that can help everyone on your team better understand the process of collecting data and how useful it can be in your work.

It is likely that your team has already collected some data to document the extent of the problems facing the community. You may have basic demographic statistics, crime and disorder reports, and stakeholder information, allowing your team to establish that a problem exists, outline who is impacted, and who might be interested in helping to find solutions. This initial data may have supported a proposal to elected officials or government agencies, or a grant application. That, however, was only the beginning of what data can provide for your work.

Understand that “data” is not just the official crime statistics in your community. Data includes all the pieces of information collected by your team or by other groups, agencies or departments to analyze your problem and then develop potential responses to that problem. This toolkit is focused primarily on raw (previously unanalyzed) data, not pre-existing statistics.

As your project moves forward, your team will enter a more intense, longer-term relationship with data. It is at this point that some project teams face early challenges. Some teams are unsure what additional data might be useful and don't pursue any information beyond the data initially collected. Others spend months trying to identify what data to collect and then more months waiting for someone to give

### Examples of Data

- Calls for service to local law enforcement
- List of attendees at a community meeting
- Pictures and addresses of blighted properties in your neighborhood
- List of addresses with code violations in the past year
- Locations of vandalism in the community

### The LISC Approach

These powerful themes run through all LISC projects.

#### DATA-DRIVEN



LISC targets crime hot spots – often streets, properties or public spaces in communities that have struggled with crime for years. Researchers are engaged in the day-to-day work, helping partners examine problems, assess evidence-based solutions, and monitor progress.

#### COMMUNITY-ORIENTED



LISC champions active roles for residents in identifying problems, selecting strategies and creating safe and healthy environments.

#### SPURS REVITALIZATION



LISC tackles problem properties, unsafe streets and parks, unemployment, transit barriers and service gaps related to crime.

#### BUILDS PARTNERSHIPS



LISC taps the resources of public, nonprofit and community leaders to bring more resources and different approaches to bear on longstanding crime challenges for lasting change.

## Collecting Data for a Safer Community, *continued*

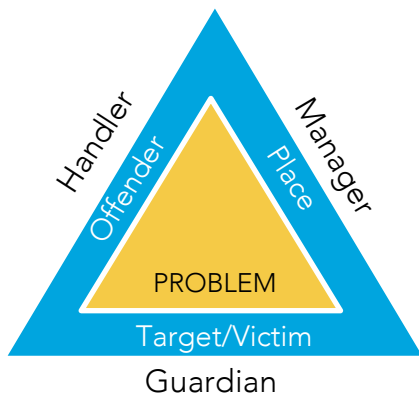
it to them. And some collect data from just one source and consider their efforts complete. These routes limit a more complete analysis of a problem and, ultimately, the development of appropriate responses. All of these challenges can be overcome with planning and foresight.

### Sources for collecting data

A more in-depth investigation of persistent crime and quality-of-life problems in the community requires more data to augment the insights of local stakeholders, inform and sharpen strategies, and make the case for change with your partners and potential supporters. But what data?

To expand your general understanding of the problem, consider the routine activities theory, which states that in order for a crime to occur, a victim and an offender must come together in time and space in the absence of capable guardians. The concept is illustrated with the problem analysis triangle, including the handlers who can control the offender, guardians who can protect the victim, and managers who can control the space.

### Problem Analysis Triangle



Source: [popcenter.org/about/?p=triangle](http://popcenter.org/about/?p=triangle)

The problem analysis triangle should serve as a reference for your data collection efforts as you think about:

- Who is involved in your problem (victims and offenders)
- Where and when those problems occur
- Who could (or does) act as a guardian, handler and manager

Your team has likely already begun to tap existing law enforcement records, such as calls for service (CFS) data, crime reports, arrest databases, traffic collision records

and field interview cards. This data should provide you with an initial understanding of each side of the problem triangle.

### Examples of Alternative Data Sources

#### Non-Traditional Criminal Justice Data

- District Attorney/Prosecutor
- Probation and/or Parole
- Jail

#### Non-Criminal Justice Agency Data

- Auditor/ Tax/ Property
- Building code
- Other Emergency Services (Ambulance and/or Fire)
- Health care

#### Community Data

- Business Groups
- Local/Citizen Groups
- Schools
- Parks and Recreation

Next, seek out more specific data to provide more context, which is the key to finding long-term solutions to the problems. For example, if your team has found that victims or offenders are juveniles, you may want to work with the local schools, parks, recreation/community centers, or other youth facilities to gather more information. If specific properties in the community appear to be chronic nuisances, the team could work with government agencies beyond law enforcement to gather more information on issues such as building code violations, property tax arrears or calls for emergency services. If overdoses are a disproportionate problem in your neighborhood, your team could work with local health clinics, emergency rooms, or needle exchange programs to gather more information—although in this case, you will need to ensure you're acting within HIPPA (Health Insurance Portability and Accountability Act of 1996) laws.

Information such as this can also alert the team to unreported crime or disorder in the community—there may be more victims of violence who have been seen in the local emergency room than cases reported to the police,

## Collecting Data for a Safer Community, *continued*

for example. Further, the information can help your team connect with additional stakeholders who may be able to provide valuable assistance in the project, from schools to the chamber of commerce to social service agencies. Although the information does not necessarily need to identify specific victims, offenders or crime incidents, that level of detail could provide more insight into how and why the problem is happening in a certain location by certain types of people to certain types of people.

### Gathering qualitative data

As we discuss data collection efforts, it is common to think of data as a list that can be summarized by numbers (i.e., quantitative data). However, data is not limited to database files. Qualitative data, such as responses to a survey, photographs or observations can help your team and partners better understand or visualize problems in the community.

In some instances, qualitative data may have already been collected by community groups or governmental agencies. If not, your project team can collect original data (also known as primary data) specifically to support your research efforts.

### Data That's Not in a System

Useful information can be gathered from:

- |   |   |
|---|---|
| <ul style="list-style-type: none"><li>Visual assessments</li><li>Environmental surveys</li><li>Maps</li><li>Time graphs</li><li>Photos/videos</li><li>Intelligence info</li><li>Police interviews</li><li>Stakeholder canvass</li></ul> | <ul style="list-style-type: none"><li>Neighborhood surveys</li><li>Business surveys</li><li>Student surveys</li><li>School personnel surveys</li><li>Parent surveys</li><li>Offender interviews</li><li>Victim interviews</li></ul> |
|---|---|

While it may feel overwhelming initially, collecting original data does not have to be a massive undertaking. It can begin by talking to a few residents or officers and recording their thoughts. Consider taking photos in the problem location(s) and documenting when and where they were taken, along with some descriptors that are relevant to the project. The process could continue to collecting names and addresses of business-owners who would be willing to participate in a business group.

This data can be collected by any partner in the project. It may be useful to match the data being collected with the right coalition partner. For example, a survey of the community about their experiences with the police will probably be more accurate if conducted by a community organization or research partner.

### Developing a data collection plan

Once your team has identified useful data sources, you will want to develop a data collection plan and establish the scope of the efforts to garner useful and actionable information. The team should consider:

- Where is the data?
- What does the data look like?
- What period of time does the data cover?
- What information is available in the data?
- Who will analyze the data?
- Do you need formal data sharing agreements?

Some pieces of data may be easy to collect while others could be more challenging. Your team can quickly go online to the U.S. Census Bureau's American Factfinder site ([factfinder.census.gov/faces/nav/jsf/pages/index.xhtml](https://factfinder.census.gov/faces/nav/jsf/pages/index.xhtml)) and download a wide variety of data about your community, from general information (e.g., demographics) to very specific data (e.g., how many people in your community own a vehicle). On the other hand, collecting a list of activities offered at the local community center can be simple—if a committee member is the one who publishes the center's schedule—or take more work if nobody on your team has a relationship with the center at all. As you consider data that helps you understand local problems, you may also want to collect and map information about the community's assets. That might involve getting input at meetings, doing online research, as well as spending time in the area of concern.

This section outlines a step-by-step process for collecting data for your project. Your team may already know the answers to some of these questions for some pieces of data. For others, you may need to research whether the relevant data is available. You may be tempted to just grab some data and start analysis. However, that route can lead to many stops and re-starts. If the team takes time now to outline the effort, you will be less likely to have to re-do work several months down the line when

## Collecting Data for a Safer Community, *continued*

you have had a chance to refine your foci (spatial, temporal, criminal or other).

### Where is the data?

Once your team has identified specific datasets that they would like to collect, you need to determine where that data is housed and who controls it. Can you confirm that the data that you would like to collect actually exists? And that the team can receive access? Does someone on your project team have a connection to that organization to request the data—or even better is in direct control of it? Is there another individual or body from whom you should seek approval before collecting the data?

### What does the data look like?

After you have identified what you need to know to access the data, you need to think about what it will look like. (Try to include someone on your project team who is knowledgeable about data formats and database terminology to be able to speak the same language as the various data providers.) In what format will you receive the data? Does the data exist in a specialized program that you can only view and not copy? Can you only receive the data in a paper format? Can the data be converted to a .csv file for transfer, but will you need to develop your own spreadsheet to review the data? Will you receive records and not a database? If so, will the records be filtered to the specific period and location (target area) that the team is examining, or will you receive records for the entire neighborhood, town/city or county? Will you receive a codebook to define the fields and special codes in a database?

Ask the data owner/manager how clean the data is and if there are any general quirks that you should be aware of when your team begins to analyze it. Many datasets have missing cases or need to be cleaned to standardize cases for analysis. Special codes (with no codebook), manually entered data that is not regularly edited for errors or irregularities (leading to difficulty with searches, summaries and analyses), and multiple unique addresses for individual properties are among the more common issues that may appear in a database.

Asking in advance about cleanliness and quirks will give a data analyst a sense of how long they will need to spend working on the data before it can be analyzed. In some

instances, data may be so dirty that it may not provide a reliable estimate of the information it was meant to describe. For example, if you are examining pickpocket reports to learn more about the victims targeted in your project area and you discover that for your reports, the data file is missing age, gender and race for 150 reports, then the analysis will not be reliable and you should seek alternative ways of gathering the information.

### What is the timing of your data?

To conduct a thorough analysis of data, it is useful to have both historic and current data. Historic data, particularly related to crime and disorder, is useful to establish long-term trends in a community. For example, information about the past five years of crimes in the target area will allow you to identify if particular crimes, offenders or places have been causing long-term problems. That would establish that there is a need for assistance in the community and help your team hone in on particular people, places or activities.

In contrast, current data is actionable data. You may know from historic data that there has been a drug hot spot in your community. Current data will help you to identify the specific corners, alleys or properties that are attracting drug buyers and sellers today. Data on crime activity should be updated regularly during the project (monthly, if possible), to account for changes in activity of victims, offenders and places. The project team should work with the local law enforcement partner to set up a schedule for regular data transfers. If you are lucky, you may be able to receive calls for service data with only a few days lag.

Data collected from other sources may not be as time-sensitive as law enforcement data, but it is still important to be aware of its age. While other factors are at play, timing can have a great impact on your analysis and your response. Is the list of local businesses in your community updated annually or is it five years old—and if so, how much has changed? Is your schedule of community center events from last winter, with an unusually limited set of activities due to snow and holidays? If your map of problem locations for youth was collected during the school year, but your prevention/safety activities will be implemented over summer break, you may not reach your target group.

## Collecting Data for a Safer Community, *continued*

**What information is available in the data?** Once you have established that the data exists, you have a sense of what its format is, and you know the timing available, you'll also want to determine what fields are available to you and if you find them useful. For example, a CFS file may document the date and time when a call to the police was received, but does the related report provide information about when the crime actually occurred? If you are reviewing theft from auto reports in your community, do the reports include detailed listings of the items stolen, where the vehicle was located (on the street, in a driveway, in a garage), and whether or not it was secured? These details could help you learn more about the actions and motives of your offenders, but only if it is included in the reports file.

Property-related data may come from your local auditor/tax collector, regional planning organization, or community development corporation (CDC) partners. It can indicate whether a property is residential or commercial and even specific details about the type of property (e.g., 10-29 unit rental housing, vacant land, gas station, convenience store). These files can also provide information about property owners or management companies that may be important to your team's efforts. However, the files may also contain dozens of coded fields that are not necessary for the project. Be sure to learn what fields are available, to determine which ones are relevant to your work, and only request useful fields in order to save file space and time in the analysis.

### Who Collected the Data and Why?

Be conscious of who collected certain datasets and why. Different groups will have different priorities in collecting information. Even within an organization, different datasets may be used to tell different stories. It is important to understand why a certain set of data was collected and what information it can and cannot provide.

This is particularly important is when examining community-generated calls for service and crime incidents, as compared to officer-generated calls. If your community has a drug problem and your only source of data is from arrests, it may not tell the full story of when, where and who is involved in drug offenses. In this case, additional data collection through community and environmental surveys and observations could be useful to better understand the full context of the problem.

If the data will be used for mapping and spatial analysis, most project teams will also be interested to learn about how address information is documented and formatted. You should identify if the address files include latitude/longitude (or other spatial) coordinates to map the data. If not, the address data will need to be geocoded to determine the latitude/longitude.

### Who will analyze the data?

The answers to the questions above only matter if your team has a person (or persons) capable of examining and manipulating the data, comprehending the findings and able to share the result with others. This person will also need to have access to the tools necessary to conduct the analysis. Many analysis efforts are disrupted when someone plans to take a class to learn a particular analytical technique, only to run into problems with scheduling, technology or ability. And remember that statistical analyses and maps are only useful for the project if there is someone who can explain what the results mean (in plain language) to the project team and other stakeholders.

Project teams should avoid working in silos, however, where anything having to do with data is only handled by the research team. Teams will learn that all parties to the project can have valuable ideas about what data to collect and how to analyze it to best understand the entire problem. Further, in some instances, the research partner may be now well connected in the community to gather data. Both community and law enforcement partners may be surprised by the relationships that they have that could help in identifying and collecting data.

### Do you need a formal data-sharing agreement?

The short answer is "it depends." Each situation is a mix of the relationship of the partnership, the type of data being shared and how it will be used. The steps noted above can be carried out informally among your team and the organizations that house the data. Even in this situation, however, you will want to have some sort of documentation that notes that "party A" will provide "X set of data" to "party B" by an agreed-upon date. This ensures that months are not lost while "A" is waiting for "B" to come to their office to collect the data, when, at the same time, "B" is expecting "A" to deliver the data.

The project team may also wish to develop formal data-sharing agreements to ensure that all parties are in agreement about what is being handed over, the

## Collecting Data for a Safer Community, *continued*

timeline and who will be conducting specific activities. Such agreements may be required among government agencies or when sharing sensitive information. A data-sharing agreement may also be used to document what should be done with the data after the project is completed. (Examples of data sharing agreements can be found at [tinyurl.com/safetydatashare](https://tinyurl.com/safetydatashare).)

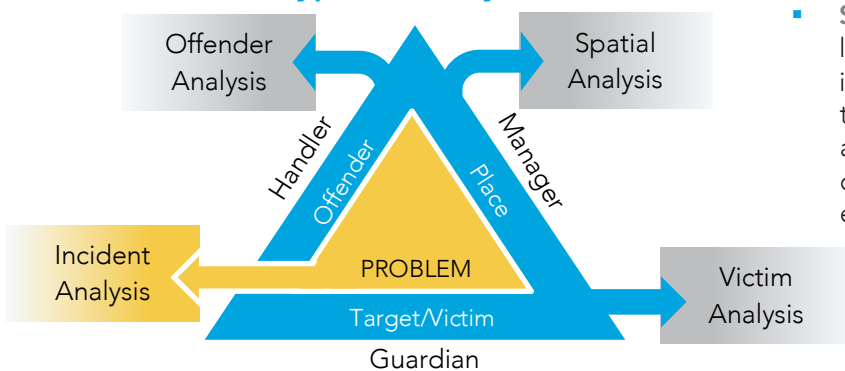
### Developing a data analysis plan

Once the team has collected data, then what? This document focuses on data collection efforts, stopping short of delving into the wide variety of analyses that can be undertaken to examine problems. However, it is useful to have a basic plan for analysis even before beginning data collection, to ensure that the proper information is available to reach your goals.

If you are using the problem analysis triangle to guide your work (highly recommended!), then your team will aim to analyze at least four different aspects of crime in the community: incidents of crime and disorder, who were the offenders, who were the victims, and the places (and corresponding times) where the incidents occur.

- **Incident analysis** includes examining the specifics around the incidents, the temporal aspects of the incidents (day, time, season), and any unique aspects to the incidents.
- **Victim analysis** should answer questions around the typical victims and targets (e.g., gender, age or activities leading up to the victimization).
- **Offender analysis** includes analyzing data about the suspected offenders, such as gender, age, school attendance, involvement in the criminal justice system or other personal attributes.
- **Spatial analysis** examines all of the various locations associated with the problem to determine if relationships exist among places. In addition to the actual crime location(s), this analysis may also examine places associated with victims and offenders, the selling or disposal of property or evidence, along with community assets or liabilities.

### Types of Analysis



As your team considers what it wants to know more about and which type of analysis will provide those answers, keep in mind that the data should really drive the specifics of what is being analyzed. With law enforcement data and data from other sources, your team can create and maintain strategies that have a real and sustained impact on public safety in your community.

### Resources and References

- Boba, R. (2003). *Problem analysis in policing*. Washington, DC: Police Foundation. [popcenter.org/library/reading/pdfs/problemanalysinipolicing.pdf](https://popcenter.org/library/reading/pdfs/problemanalysinipolicing.pdf)
- Clarke, R.V. and Eck, J.E. (2005). "Step 21: Collect your own data." *Crime Analysis for Problem Solvers in 60 Small Steps*. Washington, DC: U.S. Department of Justice, COPS Office. [popcenter.org/learning/60steps/index.cfm?stepNum=21](https://popcenter.org/learning/60steps/index.cfm?stepNum=21)
- Cohen, L. E., & Felson, M. (1979). "Social change and crime rate trends: A routine activity approach." *American Sociological Review*, 44, 588-608.
- Schmerler, K. and Velasco, M. (2002). "Primary data collection: A problem-solving necessity." *Advanced Crime Mapping Topics* (pp. 83-88). Denver, CO: National Law Enforcement & Corrections Technology Center, University of Denver. [justnet.org/pdf/cmap\\_adv\\_topics\\_symposium.pdf](https://justnet.org/pdf/cmap_adv_topics_symposium.pdf)
- Scott, M.S. (2016). *Identifying and defining police problems*. Washington, DC: U.S. Department of Justice, COPS Office. [popcenter.org/tools/id\\_define\\_policing\\_problems/index.cfm](https://popcenter.org/tools/id_define_policing_problems/index.cfm)
- U.S. Census Bureau (2018). American FactFinder. [factfinder.census.gov/faces/nav/jsf/pages/index.xhtml](https://factfinder.census.gov/faces/nav/jsf/pages/index.xhtml)

## Collecting Data for a Safer Community, *continued*

### Appendix A: Common data challenges and tips for overcoming them

Data collection efforts are rarely without problems. In the table below, we highlight a few small and not-so-small data collection challenges that we have witnessed over many years of working on problem-solving efforts. This list is by no means exhaustive, but it will give your team a sense of some of the things that have stumped other project teams in order for your team to avoid them (or at least be better prepared). Some teams have spent anywhere from hours to months fixing these problems—once they were discovered. Other projects have been completely derailed because a problem was not apparent at the onset of the project.

Challenge	Strategy
Law enforcement partner was transitioning from one records managements system (RMS) to another during the project. Data in the old system was updated for several months after the new system was introduced as cases were updated and then closed. Multiple bugs in the new system led to a delay of several months in developing profile of victim and offenders in target area.	Law enforcement partner and research partner worked together to develop complete picture of old data. Research partner ultimately worked directly with RMS development team to identify and fix bugs as they were identified.
Data provided by law enforcement partner did not have consistent labeling of numbered streets. For example, databases included both “6 <sup>th</sup> St” and “Sixth St” as valid addresses for the same location. In some instances, “Sixth Street” was also used.	Once aware of the problem, the data analyst was able to clean the dataset and conform addresses over the course of a few days. GIS could also be used to look at repeat locations if latitude/longitude or other x/y identifiers were also included in the data.
This problem is widespread among databases where individuals can enter unique data. Be aware of the potential for misspellings or variations in abbreviations.	This issue should be a consideration when cross-referencing datasets, such as fire/EMS and police. Different records systems may record the same address in a different manner.
Once a BCJI grant began, the researcher requested data from the law enforcement partner, who suggested the team use data publicly available on the Internet or submit a FOIA request.	The research partner submitted the FOIA request, but also worked with LISC staff and the entire project team to ensure that future data collection requests would be handled in a less formal, more detailed and quicker manner.
The project proposal was initially written by the community partner. When the research partner was assigned and met with local law enforcement, they discovered that crime data was not available at the address level and could only be provided by zip code.	The project team embarked on a lengthy qualitative data collection effort during the first year of the project. Surveys and focus groups were conducted among residents, youth, law enforcement and other stakeholders to gain a sense of the problems and where the most assistance was needed.
The research partner discovered that in the target area, there was an exact match in the number of “burglary in store” and “criminal trespass” offenses in commercial properties. This was consistent during any specified period of time, which seemed to be an unlikely coincidence.	After talking to law enforcement, it was discovered that every store burglary was also recorded as a criminal trespass incident. The codes were thereafter combined into one, so that every incident was not counted as two separate crimes. This fix was completed in a few hours and the research partner made a note about manually combining these incident codes in future analyses.
Law enforcement provided arrest data with addresses and latitude/longitude. After preliminary analysis, the research partner realizes that the points on the map do not match the addresses in the data.	The research partner discussed the issue with the law enforcement agency and learned that the latitude/longitude data identified the call address, while the address field was the actual arrest address. Research partner re-geo-coded the addresses.

As you can see, some of these challenges can be overcome simply by talking about data collection efforts amongst the entire project team. This includes not only the project leaders, but also the individuals who are responsible for the data (and potentially more familiar with the quirks).

### Appendix B: Team activity

At the start of their problem-solving efforts, project teams can use this activity to learn more about how basic data is collected and how it can be analyzed by the team. It can help everyone understand what data is available and the pace of data collection and analysis. It also encourages the team to think about all three sides of the problem analysis triangle.

#### Data Activity

##### How easy is it to collect your crime data and analyze it?

Collect the most recent three full months of crime report data and calls for service (CFS) data that are available for one crime problem that you will be focusing on in your community. Examples could be residential burglary, drug sales or disorder. You are not being asked for one code, but one type of crime. This may include multiple codes or labels in your records managements system or Computer Aided Dispatch (CAD) system, which typically houses CFS data. Once you have the data, meet to discuss and analyze, asking these questions:

1. What is the most recent data available?
2. What different crime types or call codes are included in the search?
3. In what format is the data and what descriptive fields are available?
4. What geographic data are included in your files? (e.g. address, X/Y, beat, district, neighborhood)?
5. Provide a count of CFS for this crime for your entire city/town for the month. Provide a count of CFS for this crime for your target area (does not need to be a hot spot, but can be if that has been identified).
6. Provide a count of reports for this crime for your entire city/town for the month. Provide a count of reports for this crime for your target area.
7. Provide a table that identifies the race, gender and ages of victims for the cases identified in the target area, based on data from the Reports file. How many cases are missing victim data?
8. Are arrest data linked to your Report files? For how many cases? Can you provide age, race and gender of arrestees?
9. Based on these findings, identify three organizations that you might contact to further develop your understanding of the problem.
10. Identify one other source of data that is available for each side of the problem analysis triangle: victims, offenders, places.

As the team works through the steps above, everyone should refer to the many questions that were discussed in the section, "Developing a Data Collection Plan," including:

- How long did it take to pull the datasets?
- Who is the best contact person to get this data?
- How recent is your data?
- Is calls for service data more recent than crime report data?
- As you consider the codes or types of incidents that are included in this data, does it seem like the report might be over- inclusive or under-inclusive?
- How complete is the data?
- How many missing cases do you have for each question? For example, how many cases do not include victim information for crimes that should have victims?
- How does your target area data compare to the data for the entire city/town?
- Are there any obvious problems with the datasets? For example, are entire columns missing information?
- What are the next steps that the team should undertake?

By conducting this activity and reviewing the results as a team, everyone will have a chance to learn more about the data analysis process. Knowing how these materials are produced can give team members ideas for other information they would like to know, as well as the options and limitations of data and its analysis.

The Local Initiatives Support Corporation (LISC) is the national technical assistance provider for BCJI, working in cooperation with the Bureau of Justice Assistance. LISC appreciates Kathleen Gallagher and Julie Wartell for their partnership to produce this paper.

This paper was supported by Grant No. 2012-AJ-BX-K046 awarded by the Bureau of Justice Assistance to LISC. The Bureau of Justice Assistance is a component of the Office of Justice Programs, which also includes the Bureau of Justice Statistics, the National Institute of Justice, the Office of Juvenile Justice and Delinquency Prevention, the Office for Victims of Crime, and the Office of Sex Offender Sentencing, Monitoring, Apprehending, Registering, and Tracking. Points of view or opinions in this document are those of the author and do not necessarily represent the official position or policies of the U.S. Department of Justice.