



THE CRIMINAL JUSTICE EVALUATION FRAMEWORK (CJEF)

The Criminal Justice Evaluation Framework (CJEF) comprises a series of documents designed to introduce people who are inexperienced in evaluation to the various methodologies available, so that they can determine the most appropriate approach for their particular evaluation. It is also designed to provide guidance to government departments and external agencies on the standard expected across the Queensland Government of evaluations conducted in a criminal justice context. While it is acknowledged that evaluating programs, policies, initiatives, interventions and operations that exist within real-world contexts requires flexibility, patience, and the capacity to communicate effectively across multiple disciplines and with persons inexperienced in research methodologies, high standards of research in all criminal justice evaluations are vital.

INTRODUCTION

Criminal justice agencies are increasingly required to demonstrate the effectiveness of crime prevention, diversion and rehabilitation initiatives, as well as new service delivery models. Evidence of program effectiveness is critical to ensure that programs are achieving their goals and not having unintended consequences. Such evaluations also inform resource allocation decisions and the distribution of limited funds across a highly competitive criminal justice sector. Yet few policy advisors and program coordinators have the time to develop significant expertise in research methods and design. This sometimes means that well-intentioned evaluations become methodologically flawed, making it difficult to meet Government requirements to provide evidence of a program's effectiveness.

Fortunately, policy advisors and program coordinators do not require a detailed knowledge of research methods and design in order to support improved evaluations and assist in the process of interpreting and critiquing program outcomes. Simply learning the basic principles of evaluation can help avoid costly mistakes and better demonstrate the outcomes of interventions.

WHAT IS AN EVALUATION?

Evaluation is the systematic collection and analysis of information to make judgements about the effectiveness, efficiency, and appropriateness of a program or initiative.

The principles of evaluation can be applied to many contexts (e.g. policies, programs, initiatives, interventions and operations; hereafter all of these contexts are referred to inclusively as 'programs'). A program typically contains a series of actions or processes designed to produce some level of measurable change or outcome - for example, a mentoring program may match at-risk youth with mentors; support the relationship over time to improve the number in education and employment; and reduce the numbers who have contact with the criminal justice system. Alternatively, police operations may be changed through the introduction of hot spot policing where their activities are focused on high-crime locations to improve community safety and reduce the number of crimes in a particular location.

Evaluation is a dynamic process that assists with the ongoing development and adaptation of programs to better suit the context in which they operate. Therefore, it is of benefit to policy advisors and program coordinators to incorporate an evaluation strategy in the early stages of program planning.

This document introduces process evaluation and outlines the standards expected of this type of research.

WHAT IS A PROCESS EVALUATION?

Process evaluations examine whether a particular program was implemented and operates as planned (Rossi, Lipsey & Freeman, 2004; Saunders, Evans & Joshi, 2005). Process evaluations are not intended to determine whether a program achieved what it set out to achieve. For that you need to conduct an outcome evaluation.¹

Process evaluations establish the extent to which a program suited the context for which it was designed, was appropriately implemented, and/or reached those persons for whom it would provide the most benefit (Saunders et al., 2005). Once problems in program implementation have been identified it is possible to feed this information back into the implementation process and thereby improve the program. Where a process evaluation is conducted together with an outcome evaluation, the process evaluation can assist to determine *why* a program was effective or not in achieving its intended outcomes.

HOW TO CONDUCT A PROCESS EVALUATION

A process evaluation is conducted in four steps:

1. Identify how the program was intended to be implemented and to operate.
2. Identify how the program was actually implemented and actually operates.
3. Analyse data to determine whether the program operates as intended, and if not, why not.
4. Develop recommendations and communicate findings.

STEP 1: IDENTIFY HOW THE PROGRAM WAS INTENDED TO BE IMPLEMENTED AND TO OPERATE

A process evaluation begins by identifying the way in which the program was intended to be implemented and to operate. With respect to the way the program was intended to be implemented, you will need to clearly establish:

- the intended recipients of the program
- how the program was intended to be accessed and what processes were to be put in place to facilitate access
- how the program was intended to be resourced
- those factors expected to impede implementation and how they were to be addressed
- those factors expected to facilitate implementation and how they were to be enhanced
- the precise steps to be undertaken when rolling out the program.

To assess how the program was intended to operate, you will need to clearly establish:

- the issue/s the program is intended to address
- the activities intended to be undertaken as part of the program
- the intended organisational structures and staffing patterns for the program
- the products, goods and services intended to be produced by, or as a result of, the program
- the expected outcomes of participation in the program
- the intended relationships between inputs,² outputs,³ and outcomes⁴ (i.e., how and why each input is expected to produce a particular outcome)

¹ For further information regarding this type of evaluation see *Criminal Justice Evaluation Framework (CJEF): Conducting effective outcome evaluations*.

² An input is any resource used to produce a program output (e.g., program facilitator, funding support for the program, transport support such as a bus pass).

³ An output is any good or service produced by a program for the purpose of achieving an outcome (e.g., case management plan, participant accessing support services (because of a bus pass)).

- the context in which the program was intended to operate.

It is important that you operationalise these intentions in a way that is clear, concrete and measurable. That is, you need to precisely determine how you will know when a program is operating in the way in which it was intended and when it is not. For example, clear and concrete indicators of the extent to which a program delivers support services to youth in remote communities at risk of entering the criminal justice system may include:

- the precise locations in which the program was intended to be implemented and the order in which it was to be implemented across locations
- the number of youth expected to be referred to the program
- the proportion of youth entering the program expected to be referred to support services
- the proportion of youth entering the program not expected to be referred to support services and the precise reasons why
- the number and quality of support services thought to be available and accessible to program participants
- the expected cost of each referral.

These measures are often referred to as *administrative standards* and may be established based on similar or past initiatives, on the basis of the expertise of program developers and stakeholders, or on established scientific theory (Rossi et al., 2004). At other times a program may be expected to meet established legal, ethical or professional standards (e.g., UN Charter of Human Rights) and it is these standards which form the grounds for determining whether a program operates or was implemented as intended. By quantifying and clarifying the intentions surrounding a program you will be better able to determine whether the way in which a program actually operates matches the expectations of program developers.

In order to best quantify or clarify an intended activity, service or context, you should precisely define its components in terms of its smallest part. Defining an activity in terms of its smallest part will enable you to ascertain whether there are differences in the way that a task is performed, which in turn can help to explain why a program is more or less successful in achieving its intended outcomes. For example, when referring a young offender to support services a program facilitator may be expected to:

- conduct a needs assessment
- identify services congruent with the offender's needs
- complete a referral form
- follow-up the young person's attendance at the referral agency.

An evaluator may find that all program staff conduct a needs assessment; however, the way in which they perform this activity may differ across locations and from the procedures put in place during the program's development. This may explain any observed lack of fit between a young person's needs and the service to which they are referred. It may also be used to inform recommendations for staff training or to account for any lack of program effectiveness observed in an outcome evaluation.

That said, it is sometimes inappropriate to define an activity, service or context in terms of its smallest part. This is because the components of the activity, service or context may be different for each participant and therefore could not be compared across participants. For example, in the program described above, the referral form completed on behalf of each young person may differ depending on the service to which they are referred (e.g., support services for drug abuse versus homelessness). It would be inappropriate to compare the extent to which information specific to each service (e.g., detailed information regarding drug

⁴ An outcome is the subsequent effect of a program, typically as a consequence of program outputs (e.g., reduction in reoffending behaviour, reduced drug use).

use) was provided on forms for different agencies. Instead, evaluators may like to examine broader commonalities and differences between forms (e.g., the inclusion of offence history, family background and results of the needs assessment) and whether the inclusion or exclusion of this broad information facilitated the referral process.

The methods used to collect data about the way that a program was intended to operate are similar to those used to collect data about the way that a program actually operates (described in more detail below). For example, each of the following can be used to establish the way that a program was intended to operate or actually operates:

- interviews or surveys with program developers, key stakeholders and participants
- content analysis of program guides and policies
- analysis of documents which describe the way that a program was designed (e.g., documented implementation plan) as well as descriptions of other programs or theoretical positions relied on by program developers during program development
- examination of parliamentary debates, ministerial statements, and explanatory notes in the case of changes introduced through legislative reform.

STEP 2. IDENTIFY HOW THE PROGRAM WAS ACTUALLY IMPLEMENTED AND ACTUALLY OPERATES

Once you have identified the way that the program was intended to be implemented and operate, you are ready to investigate how the program was actually implemented and operates. Again it is important that you rely on clear, concrete and reliable indicators.

There are several methods that can be used to collect this information. Using a combination of approaches is recommended, as this allows for cross validation of information and reduces the likelihood that inaccurate or inappropriate conclusions are made (Bouffard, Taxman & Sliverman, 2003). That said, try to avoid selecting too many indicators as this may make the data collection process onerous, expensive and/or lacking in relevance to your research goals. In addition to identifying the type of data you will collect, it is also important to consider the factors likely to constrain or facilitate data collection. These include:

- the source of your data (i.e., where it will come from and the ease with which it can be obtained)
- the quality of the data available (i.e., is it complete, reliable and valid)
- how it will be collected (e.g., surveys or departmental records)
- who will be responsible for collecting the data
- the timeframes for data collection (e.g., will data be collected weekly, fortnightly or monthly).

Given the known difficulties associated with collecting criminal justice data,⁵ identifying how your data can be collected may assist in limiting and refining your collection methods to those which are likely to be most informative, reliable and accessible for your context.

A brief description of data collection methods which are typically used in process evaluations is provided below. The advantages and disadvantages associated with each are listed in Table 1 at the end of this section.

Stakeholder interviews

Those people who are responsible for the day to day running of a program typically have the greatest insight about its implementation and operation. They are therefore a valuable source of information. It is important, however, that evaluators avoid conducting stakeholder interviews in an ad hoc fashion and instead rely on

⁵ For more information see: *Accessing Queensland Government Data for Criminal Justice Research: A guide for academics and external agencies*

structured or semi-structured methods. In structured interviews participants are asked exactly the same questions in exactly the same order. In semi-structured interviews, participants are asked exactly the same questions in exactly the same order and given the opportunity to elaborate on issues raised by their responses through additional questions developed by the interviewer in situ. The benefit of conducting structured or semi-structured interviews over ad hoc interviews is that they allow evaluators to compare participants' responses to the same questions. This means you will be able to identify similarities and differences in participants' responses.

Content analysis of program materials

Content analysis involves reviewing written materials in a systematic way to identify common themes and to categorise the nature of their content. Any documents associated with the actual or intended implementation, promotion and administration of the program are likely to be appropriate for analysis in a process evaluation. Documents may include policy manuals, pamphlets or education materials provided to participants. For a comprehensive discussion of approaches to content analysis see Bryman (2008).⁶

Monitoring processes

Evaluators may also rely on records kept by program facilitators during the course of the program. Record keeping by program facilitators is referred to as monitoring. In an ideal situation, an evaluator will work with program developers and facilitators from the outset to develop methods for efficiently and effectively monitoring program processes. This is the best way in which to ensure that the information collected by facilitators is appropriate for use in subsequent evaluations and to inform ongoing program development.

Monitoring is a term often used interchangeably with evaluation. However, the two processes are quite distinct. Monitoring is the continuous and systematic collection of information for the purpose of informing the extent of progress towards pre-determined goals (Rossi et al., 2004). For example:

- Queensland Corrective Services may record the number of Indigenous prisoners who participate in a given program for the purpose of monitoring the extent to which a goal of 80% participation is achieved.
- Queensland Treasury may monitor spending on a particular project to ensure the limits of funding are not exceeded.

Unlike data gathered for the purpose of evaluation, data gathered through the monitoring process is not typically used to draw conclusions as to the effectiveness of a program or to identify ways in which a program may be improved.

Records kept by program facilitators for monitoring purposes may be used by evaluators to identify what is actually delivered by a program and to whom it is delivered (i.e., program coverage and program bias). For example, service delivery measures of a drug treatment program might include the number of graduates from the program, the number of treatment sessions administered, or the number of information/promotional pamphlets given to potential participants. The extent to which persons for whom the program was intended (target population) participate reflects program coverage and program bias (Rossi et al., 2004).

An analysis of program coverage considers whether the number of persons from the target population who access the program is consistent with the intended number of target participants (Rossi et al., 2004). An analysis of program bias considers whether subgroups within a target population are more or less likely to participate (Rossi et al., 2004). Bias may be a function of self-selection (e.g., a particular subgroup is more likely to volunteer for the program) or program actions (e.g., failure to advertise the commencement of

⁶ Colorado State University provides a free guide to content analysis at <http://writing.colostate.edu/guides/research/content/> [accessed August 2011]

programs in an appropriate way or to conduct the program in an appropriate location for a specific subgroup). In order to measure the extent to which a participant population is biased towards a particular subgroup of the target population, evaluators can draw on monitoring information to determine the characteristics of those who drop out of the program or who are eligible but do not participate.

Surveys

Surveys can also be used as a means of collecting information about the way that a program was implemented or operates. For example, participant surveys may be used to identify participant characteristics and the extent to which there is implementation bias among the individuals engaged by the program. Community surveys may be used to determine why potential candidates were not or did not engage in the program. Stakeholder surveys can be used to collect anonymous information regarding the way that a program was implemented or operates.

When conducting survey research it is important that the sample is appropriately representative of the actual participant population, or, where the parameters of the actual participant population is unknown (e.g., a television-based education program which is broadcast nation-wide), representative of the eligible participant population.⁷

Finally, it is important that the survey questions relate specifically to the goals of the program and the extent to which they were or were not achieved, and why. This ensures that the information collected by surveys can be used effectively to draw conclusions from the process evaluation.

Systematic social observation

Systematic social observation is a data collection technique used to systematically identify the characteristics of a particular social environment (Bouffard et al., 2003). Data is collected through observations structured around pre-determined categories and specific research questions. It is important that observations are recorded according to structured, replicable procedures as this allows for comparison across time and locations. Observations are typically made by research officers who are extensively trained in the purpose and process of observation. Systematic observations will increase the value of a process evaluation as they allow evaluators to use the scientific method to observe program implementation first hand and to compare results with stakeholders' descriptions of program operations (Bouffard et al., 2003). It may also be worth noting that observations made consistently over a long period of time can sometimes be more informative than those made in the short term. This is because:

- participants who alter their behaviour in response to observers may be more likely to relax over time
- evaluators are exposed to more and varied situations
- evaluators may be able to follow one cohort of participants throughout their involvement in the program.

Taking significant steps to guarantee anonymity and confidentiality will also assist to discourage participants from behaving in ways other than they normally would.

⁷ For an introduction to selecting a representative sample, see *Criminal Justice Evaluation Framework (CJEF): Conducting effective outcome evaluations*.

Table 1. The advantages and disadvantages associated with different data collection techniques used in process evaluations.

Method	Advantages	Disadvantages
Stakeholder interviews	<ul style="list-style-type: none"> - unique insight into program implementation and administration 	<ul style="list-style-type: none"> - the information provided by stakeholders may be inaccurate or incomplete because of bias and the extent of information about the program to which the stakeholder is exposed
Content analysis	<ul style="list-style-type: none"> - credible, non-biased data source 	<ul style="list-style-type: none"> - this approach to analysing information can be time consuming
Monitoring processes	<ul style="list-style-type: none"> - a clear quantitative measure of service delivery 	<ul style="list-style-type: none"> - accuracy is dependent on the accuracy with which the information is collected - extracting, analysing and interpreting data can be time consuming if it was not collected for the purposes of substantiating an evaluation
Surveys	<ul style="list-style-type: none"> - inexpensive and fast to conduct - can be used with large, dispersed and/or geographically diverse populations - can be used to address many topics or to ask many questions about a single topic - allow for the same information to be collected from different groups and over time - usually highly reliable method of data collection 	<ul style="list-style-type: none"> - to make questions relevant to a variety of participants surveys often generalise across concerns which has the potential ignore subtle differences in participants' experiences - inflexible (i.e. to enable comparison over time questions cannot change) - difficulty gaining high response rates from participants - cannot control for/account for the impact of the broader context in which the program is administered on participants' responses - potential for response bias (i.e. the participants answer the question the way they think evaluators want them to) - can take time and effort to design (e.g., may require extensive pilot testing to ensure questions address the issue appropriately)
Systematic social observation	<ul style="list-style-type: none"> - data can be collected in a way that is independent of any individuals directly involved in program implementation 	<ul style="list-style-type: none"> - can be time and resource intensive

STEP 3. ANALYSE DATA

Analysis of the data should aim to do the following (Rossi et al., 2004):

- describe the way in which the program actually operates
 - who participates, who does not participate and why
 - what services are provided by the program and what outputs are produced
 - how do participants perceive/use the services provided
- compare the way in which a program operates across different sites (when a program has been implemented in more than one site)
 - demonstrate any diversity across program sites (e.g., staff, administration, and broader context)
 - identify whether it is appropriate for activities or outputs to be standardised
- describe the extent to which the program conforms with its intended design
 - identify whether a program performs its functions as intended
 - where a program does not function as intended describe why
 - determine whether it is appropriate to next conduct an outcome evaluation to determine the effectiveness of the program.

Quantitative data (i.e., numerical data) should be analysed using appropriate statistical methods and presented appropriately in text, tables or graphs. Qualitative data (i.e., descriptive accounts, observations and interviews) should be analysed (using available software packages where the quantity of qualitative data collected makes this appropriate) to identify major themes, categories, issues and incidents relevant to the evaluation. Particularly succinct or expressive quotes should be used to support or provide examples of any conclusions drawn from qualitative data.

STEP 4. MAKE RECOMMENDATIONS AND CLEARLY COMMUNICATE YOUR FINDINGS TO OTHERS

Communicating your findings with conciseness and clarity will improve the chances of your recommendations being understood and implemented. Effective communication begins by identifying who is most likely to read your report (e.g., program stakeholders, those directly involved in the program, policy and decision makers, or a combination of these) and their motivation for reading it (e.g., what will your recommendations be used for?). Knowing your audience and their needs will then help you to decide what information should be included in the report, the way in which it should be structured, and how to support the argument you are making (e.g., the program operates as intended; the program needs to be changed). For example, it may be more appropriate for your findings to be released in a series of papers, each of which targets the select needs of a unique audience, than as a single report. Alternatively, your audience may respond better to workshop-style presentations or community meetings.

While the needs of your audience should play a key role in determining the way in which you communicate your findings and to some extent the level of detail you provide, there are some key components every process evaluation report should include. These are:

- a description of the program being evaluated and the context in which it operates
- a statement containing the specific research questions addressed
- an explanation of the methods and measures used to collect the data
- the results found from the evaluation
- any limitations of the data, data collection and evaluation
- a clear explanation of the answers provided to the research questions on the basis of the data collected
- any recommendations made on the basis of the results

- the need to move the implementation of a project closer to the original design and how this might be achieved
- the need to change the design of the original program to better suit the participants or context of implementation
- the appropriateness of conducting an outcome evaluation
- description of the way in which information will be fed back to improve program implementation.

MORE QUESTIONS ABOUT EVALUATING?

If you are having trouble establishing a good evaluation framework or have any questions about evaluation please contact Criminal Justice Research, Department of the Premier and Cabinet (Ph: 32278436) or consult the references listed below.

REFERENCES

Accessing Queensland Government Data for Criminal Justice Research: A guide for academics and external agencies. Available on the Department of the Premier and Cabinet website [<http://www.premiers.qld.gov.au>]

Bouffard, J.A., Taxman, F.S., & Silverman, R. (2003). Improving process evaluations of correctional programs by using a comprehensive evaluation methodology. *Evaluation and Program Planning*, 26, 149-161.

Bryman, A. (2008). *Social research methods*. Third Edition. Oxford University Press. Oxford, UK.

Criminal Justice Evaluation Framework (CJEF): Top five tips for conducting effective outcome evaluations. Available on the Department of the Premier and Cabinet website [<http://www.premiers.qld.gov.au>].

King, J.A., Morris, L.L., & Fitz-Gibbon, C.T. (1987). *How to assess program implementation*. Sage Publications, California, USA.

Rossi, P.H., Lipsey, M.W., Freeman, H.E. (2004). *Evaluation: A systematic approach*. Seventh Edition. Sage Publications, California, USA.

Saunders, R.P., Evans, M.H., & Joshi, P. (2005). Developing a process-evaluation plan for assessing health promotion program implementation: A how-to guide. *Health Promotion Practice*, 6(2), 134-147.

Writing Guide: Content Analysis. Available on the Colorado State University website <http://writing.colostate.edu/guides/research/content/> [accessed August 2011]