REFEREED ARTICLE
Evaluation Journal of Australasia, Vol. 6 (new series), No. 2, 2006, pp. 6–15

'What if': the counterfactual in program evaluation

Rick Cummings

Dr Rick Cummings is a Senior Lecturer in educational development and evaluation in the Teaching and Learning Centre at Murdoch University, Perth, and is currently President of the AES. Email: <r.cummings@murdoch.edu.au>

How do evaluators provide the best evidence of program outcomes to stakeholders? The current focus on impact evaluation and the adjoining debate over the suitability of various designs for rigorously measuring program outcomes begs the question of what are the most suitable alternative outcomes with which to compare program outcomes. In essence this is asking what is the counterfactual, how should it be measured and what is an appropriate method of comparing the outcomes of a program and the counterfactual. Although measurement of outcomes is a major issue in evaluation, the counterfactual does not appear often in the evaluation literature, whereas it is widely used in the fields of history, political science, philosophy and psychology. This article attempts to commence the process of filling in this gap by focusing on the role and measurement of the counterfactual in program evaluation. The underlying logic of the counterfactual is explored and some approaches to its measurement are discussed. The approach is described in detail using an example from a recent evaluation study of a crime prevention program in Western Australia. The potential benefits of measuring the counterfactual are explored, particularly in relation to providing influential information to key stakeholders.

Evaluation studies may be undertaken to provide information relevant to addressing a wide range of questions. Frameworks for deciding which questions are highest priority and what information might be most useful to address these questions have been developed by a number of authors over time (see Worthen, Sanders & Fitzpatrick 1997; Stufflebeam 2001; Owen 2006). In examining the frameworks that include a focus on measuring program outcomes, it is possible to identify three key areas of information that many evaluation studies are expected to address:

- First, an accurate description of the program, including an explicit articulation of its underlying logic or theory. This is often critical if the causal assumptions embedded in the program are to be properly understood.
- Second, identification of appropriate outcomes and measurement of them using rigorous methods. One valuable approach in this area is to measure the counterfactual (i.e. what would have happened in the absence of the intervention), as this may be essential for making causal inferences about the effect of the initiative.

- Third, information other than short-term outcomes measures is increasingly sought to inform stakeholders about the sustainability of the program and/or identify possible alternatives.

This article focuses on the second of these, in particular the use of the counterfactual. This area often is not addressed in evaluation studies, usually because of the difficulty of identifying control or comparison groups and/or describing an alternative outcome of the program. The purpose of this article is to show how the counterfactual can be a very useful approach to measuring outcomes and convincing stakeholders of the utility of the evaluation findings.

Defining the counterfactual

Counterfactual inference has been a source of research and debate in a range of fields for over 40 years, commencing in history, philosophy and statistics, but expanding to psychology, political science, computer science and other disciplines. There is an interesting and complex literature on both the psychology of counterfactual thinking and its place in western critical thought (Kvart 1986, Lewis 1973). As Mandel, Hilton and Catellani (2005) point out in their recent overview of counterfactual thinking, it is an essential component of human cognition, and is a particularly important element in understanding the Western concept of causality. Interestingly, counterfactual thinking has been found to influence conceptual understanding, satisfaction levels, and emotional responses to both situations and outcomes. Counterfactual thinking also can be seen to occur as either an evaluative (comparing the factual and possible counterfactuals) or reflective (experiential perceptions of what happened or might have happened) mode of thinking. In the area of research, Lebow (2000, p. 558) argues that, 'Counterfactuals are an essential ingredient of scholarship. They help determine the research questions we deem important and the answers we find to them'. Although there is insufficient space to explore these ideas here, they do lay the basis for examining the role and utility of counterfactual analysis in program evaluation.

In the field of evaluation, Mohr (1995) points out that the use of the counterfactual is essential in impact evaluation, as it provides the alternative against which the program's impact can be measured. In its broadest sense, the counterfactual is an estimate (either quantitatively or qualitatively) of the circumstances that would have prevailed had a policy or program not been introduced. By comparing counterfactual outcomes with outcomes measured for the target group(s) subject to the policy or program, causality or attribution can be established, because it is assumed the changes in the target are due to the intervention of the policy/program and not other factors. In theory, causality can be attributed to the intervention because it is assumed there are no systematic differences between the program group and a 'true' counterfactual group, except for the fact that the program group has been exposed to the intervention. In quantitative studies, differences in average outcomes between the program group and the 'true' counterfactual group therefore represent an unbiased measure of the program's impact.

In the context of program evaluation, the counterfactual is defined as what might have occurred if the particular program being evaluated was not in place or an alternative program was conducted. This involves measurement of both what has happened as a result of the program as well as what might have happened in its absence. This can be done either quantitatively by measuring the positive outcomes that would have been observed among the eligible population if the program were not in place or qualitatively by the development of alternative histories, scenarios or experimental thinking.

The fundamental difficulty in using the counterfactual in impact evaluation is that whereas the outcomes of the program being evaluated are observable (although not always easily identified or measured), the outcomes in the absence of the program, at least in principle, are counterfactual and not observable. This situation requires that the evaluation design provide some basis for constructing a credible estimate of the outcomes for the counterfactual conditions. In order to identify the most suitable counterfactual, the evaluation study must be designed to account for the circumstances of the particular program being evaluated, the nature of its target population, the outcomes of interest, the data available, and the constraints on collecting new data. As a result, it is difficult to define a 'best' design for impact evaluation a priori. Rather, the task for the evaluator is to identify the best design for a particular program under the particular conditions when the evaluation is undertaken. As such counterfactual analysis in program evaluation can be seen to have three broad aims (UK Prime Minister's Office 2004):

- to establish evidence of a causal relationship between a program and the outcomes the program seeks to influence
- to account for confounding factors, additional to the influence of the program, that might lead to measured change in outcomes
- to provide estimates of the impact of the program.

Measuring the counterfactual in evaluation

Perhaps the most interesting characteristic of the counterfactual is that it is possible to construct it in either quantitative or qualitative terms, depending on the nature of the program and the outcomes it is designed to produce. Although the quantitative approach has attracted most of the attention in program evaluation, qualitative approaches can be

constructed from the use of counterfactual analysis in other disciplines, for example political science, history and sociology. The development and use of qualitative counterfactuals is a complex subject that requires a separate discussion. The remainder of this paper focuses on counterfactual analysis using quantitative data.

In writing for the UK government about policy evaluation, Purdon et al. (2001) have summarised many of the characteristics of the counterfactual in evaluation, although from a clearly quantitative perspective. They point out that a measure of the counterfactual is useful in at least the following four ways:

- when a good estimate of 'added benefit' or 'additionality' is needed for a cost–benefit analysis
- when there is a need to convince others that a program or policy is beneficial
- when there are alternative policies, program or methods of implementation that might be used and there is a need to establish which is preferable
- when there is a need to establish whether a policy or program works better among some set of alternatives.

The counterfactual is particularly relevant when the outcomes for a program are expressed in terms of change, for example in terms of *increasing* the number of persons in work, or *reducing* the numbers on low incomes. In many programs there is a strong case for assuming a particular counterfactual, for example a new service might be introduced where the main target is to encourage a take-up rate above a fixed percentage each year. In this instance, the counterfactual would be no take up in the absence of the program. In other circumstances the program outcome may be a change in the rate of some existing behaviour, for example crime rate or unemployment rate. For programs addressing these social problems, the counterfactual is the existing rate of behaviour in the absence of the program. In both of the above, it is also possible to compare the change to the change produced by another program addressing the same social issue. The example discussed later in this article fits within the former scenario—a new program in the absence of an alternative. However, Purdon et al. (2001) point out that the counterfactual will be very difficult to estimate reliably if the program impact is expected to be very small or if the program impact is smaller than variation that occurs naturally over time in the eligible population (e.g. seasonal variations in the unemployment rate).

In practice, measuring the counterfactual is a difficult task. Evaluators have available to them a variety of study designs, depending on circumstances and opportunities open to them, to estimate the 'true' counterfactual. Purdon et al. (2001) provide a useful critical summary of a number of study designs in which the counterfactual can be measured and compared with the program outcomes. Suffice it here to say that these designs can be categorised into three general types: randomised trials, matched comparison design, and before–after studies. The strengths and weaknesses of the range of possible study designs are discussed in some detail in Purdon et al. (2001), and in many texts on research methods.

Many of the practical problems of the study designs identified above can be better controlled if whole areas are divided into intervention and control groups, but for practical reasons this is usually difficult to do. Instead, many programs are implemented initially using pilot projects. These enable testing of new initiatives, or changes in existing policy or practice, in a limited number of geographical areas prior to introducing them more widely. The objective is to determine whether the new initiative gives rise to changes in the outcomes that the program seeks to alter. For example, counterfactual analysis might answer the question: Is there a direct relationship between a new initiative to cut car crime and subsequent change in the number of reported car thefts, independent of other factors influencing car theft? Counterfactual analysis explicitly acknowledges the fact that the outcomes the program attempts to influence are subject to a range of factors beyond the immediate scope of the program being studied. The pilots provide groups experiencing the intervention and the rest of the population not, providing a range of possible comparison groups, although strictly speaking seldom are these randomly assigned so they fail to qualify as control groups.

An example of this type of program implementation was recently evaluated in Perth Western Australia and provides an example of the use of the counterfactual to measure program impact.

Operation Burglary Countdown

As a result of growing concern that Perth had the highest rate of residential burglary of the Australian capital cities, the Western Australian Government launched the Burglar Beware Campaign in December 2003. This campaign operated within the State Community Safety and Crime Prevention Strategy, and was overseen by the Office of Crime Prevention in the Department of Premier and Cabinet and the State Burglary Reduction Task Force. In an attempt to test an integrated comprehensive community-based strategy to reduce residential burglary in hotspot locations, the Office of Crime Prevention provided funding of $150 000 for two Burglary Reduction Pilots to run for 12 months, from November 2003 to October 2004. The Burglary Reduction Task Force selected Bentley and Morley as the two pilot sites for this initiative to be known as 'Operation Burglary Countdown'.

The pilots were based on a partnership approach and specifically sought to make use of the rational choice theory of offender behaviour (Cornish & Clarke 1986). This was done through the

establishment of Local Management Groups in each suburb to coordinate a series of processes aimed at reducing the offender's confidence in conducting a successful burglary, making the residential burglary more difficult to commit, and making apprehension and conviction more likely. As can be seen in the model describing Operation Burglary Countdown in Figure 1, the focus is on using a range of interventions in a locally managed and integrated program.

Operation Burglary Countdown is underpinned by the following four principles:

1 Elevating the offence across government agencies and in the community so that what was a community priority but not an agency priority is now also an agency priority, for example the Education Department targeting secondary school students on the street through stronger truancy and diversion programs.

2 Establishing joint ownership of the initiative by the key stakeholders so that each stakeholder not only participates but contributes to achieve the initiative's objectives, for example using council workers in an 'Eyes on the Street' crime detection program.

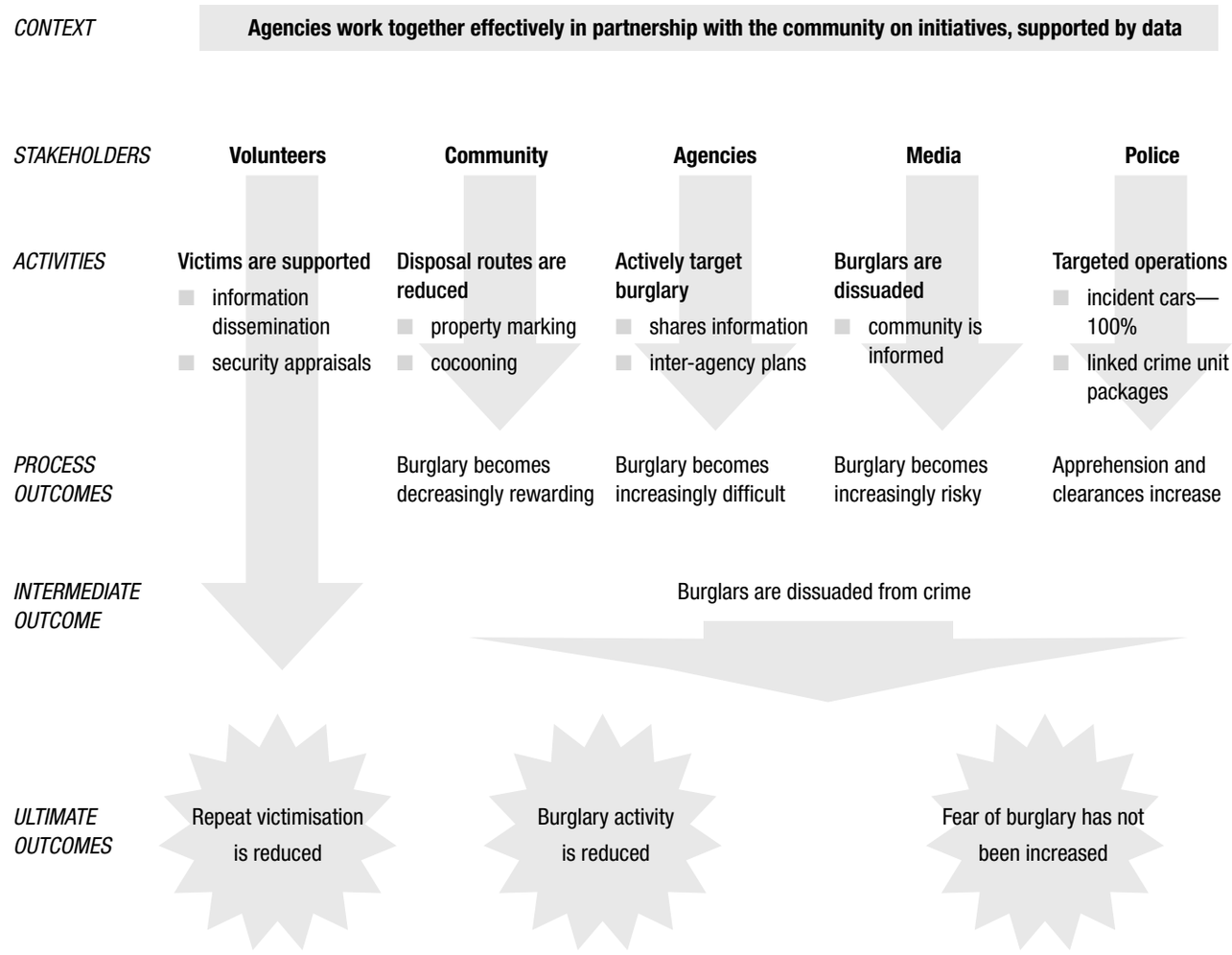
3 Implementing strategic change through building interagency relationships, aimed at improving the physical environment, reducing fear, targeted police operations and focusing on offender behaviour.

4 Building on success by careful monitoring of outcomes, marketing of achievements, and motivation of local key individuals, for example a strong local media campaign highlighting achievements in the pilot areas.

The pilot program of Operation Burglary Countdown was the subject of an independent evaluation study conducted by Estill and Associates in conjunction with Murdoch University. The evaluation study had two major purposes:

- to provide information on the implementation process and short-term outcomes for the two pilot programs

FIGURE 1: BURGLARY REDUCTION PILOT PROGRAM: PROGRAM LOGIC MODEL



- to provide an evaluation model for possible further expansion of the program.

In addressing the first purpose, the evaluation team developed a multi-method evaluation study design, as is appropriate for a program with a range of interventions and diverse stakeholders (English, Straton & Cummings 2002; Rosenbaum 2002). The evaluation study focused on three areas recommended in this literature and developed a total of 10 key evaluation questions as listed in Figure 2.

The program logic model in Figure 1 was developed at the commencement of the evaluation study and was used to guide the design of the study and the reporting of the results. A full description of the evaluation study and findings can be found in Cummings (2005).

The three ultimate outcomes identified in the program logic model that align with the overall outcomes listed above, were each measured in the evaluation study as follows:

- *Community attitudes to residential burglary.* Changes in community attitudes were measured through a pre- and post-program survey of a sample of residents in each community. The results show a mixed result with no consistent pattern of attitude change. About 65% of those who had a burglary in the past 12 months were satisfied with the support received from police and believed police provided handled it quite well or very well. However, about 75% of residents felt quite safe or very safe in their homes and this didn't change over the period of the program.
- *Level of repeat victimisation.* Research has shown consistently that victims of residential burglaries are often targeted for a repeat burglary in a short time after the original offence (Henderson 2002). Operation Burglary Countdown had a clear objective to reduce the level of repeat victimisation as one component in the overall program, and the home security audits and neighbourhood cocooning initiatives are directed specifically at this outcome. In the 12-month period preceding Operation Burglary Countdown (November 2002 to October 2003), there were 83 victims who were burgled more than once in Bentley and 26 in Morley. During the period of the program (November 2003 to October 2004), these numbers had declined to 42 in Bentley and 11 in Morley, or 49% and 58% respectively. Police records show there was not a corresponding decline in repeat burglaries in the wider Perth metropolitan area.
- *Reduction in residential burglary.* The number of residential burglaries was measured using the records of reported crimes collected by the Western Australia Police Services[1]. The results for the target suburbs, surrounding areas and the whole Perth metropolitan area are summarised in Table 1. It can be seen that the number of residential burglaries decreased in both pilot areas. However, the number of residential burglaries also declined in the surrounding areas and when compared to the decline in metropolitan Perth, which also dropped over this time, it is not obvious whether the program had a net impact on the burglary rates in the two suburbs. An excellent method for clarifying this issue is to use a counterfactual analysis, to compare what actually took place within the target areas with what might have taken place if the program had not been implemented. This was undertaken and is described below and summarised in Table 2.

FIGURE 2: KEY EVALUATION QUESTIONS

Program performance	
1	What is the value of 'partnerships' with the community?
2	What is the value of 'partnerships' between agencies?
3	How much do these initiatives cost (essential/other) (dollars/staff)? For what results?
4	Is the data collection/sharing/utilisation appropriate and effective?
Victim action and response	
5	Are victims increasing the level of security in their homes?
6	Are rates of residential burglary reporting increasing?
Overall outcomes	
7	Have community attitudes regarding residential burglary changed?
8	Has residential burglary been reduced in the pilot area?
9	Has residential burglary been displaced (geographically/nature)?
10	Has repeat victimisation been reduced?

TABLE 1: SUMMARY OF CHANGES IN RESIDENTIAL BURGLARIES AND ESTIMATED COMMUNITY SAVINGS

	Bentley Bentley	Bentley surrounding areas	Morley	Morley surrounding areas	Metropolitan Perth
Previous year 2003	647	1202	364	1029	30 818
Program year 2004	355	861	276	817	22 897
Gross change	-292	-341	-88	-212	-7 922
Percentage change	-45%	-28%	-24%	-21%	-26%

The counterfactual analysis

There is considerable research on measuring the outcomes of crime prevention programs but little agreement on either the best outcomes measures or how they should be analysed (Johnson et al. 2004: Rosenbaum 2002). However, there is general agreement that, where possible, evaluation studies should move beyond the simple approach of reporting changes in rates of reported offences, to trying to measure the counterfactual (i.e. what would have happened in the absence of the intervention) and if possible to conduct a cost–benefit analysis. It is this level of analysis that is most useful in determining which changes to crime rates can be attributed to specific interventions and estimating the net value of the resulting benefits, if any.

Johnson et al. (2004) point out that finding a comparison area that is similar to the target area can be very difficult in the evaluation of crime prevention programs. It is suggested that surrounding areas are often as close as one is likely to get, an effect known as autocorrelation. However, these areas are open to two contrasting and confounding influences: first, there may be higher than expected crime rates in the neighbouring areas due to spatial displacement of crime from the target area, or second, there may be a reduction in crime in surrounding areas due to the diffusion of the initiative to the surrounding areas. In the evaluation study of Operation Burglary Countdown, the surrounding areas for Morley and Bentley comprised 5 and 7 suburbs respectively. An initial analysis of the changes in residential burglary rates in the surrounding areas indicated that diffusion occurred in Bentley and its surrounding areas whereas a small degree of displacement occurred in Morley and its surrounds. To avoid the influence of either diffusion or displacement, the comparison area used for this analysis is the entire Perth metropolitan region excluding the pilot areas and the surrounding areas. This area is sufficiently large not to be unduly influenced by local crime reduction initiatives or influences.

Johnson et al. (2004) have developed a method of calculating the predicted number of offences prevented by a particular intervention, which is of considerable assistance in providing a method by which evaluators can measure the counterfactual. Their approach is based on a comparison between the actual crime rate in the action area before and after the intervention and the rate in a comparison area where there was no intervention. This is termed a 'differences in differences approach' (Purdon et al. 2001). Although there are known problems in identifying appropriate measures of the rate of crime and comparable comparison areas, the approach does provide a convenient and fairly simple process by which this analysis can be done. An additional benefit of this approach is that by reporting the outcome in number of burglaries prevented it provides a means of quantifying the benefits of a specific intervention, and where the unit cost of the offence can be estimated, this benefit can be quantified in financial terms, greatly assisting in cost–benefit analysis.

In its simplest form, Johnson's approach involves subtracting the observed number of crimes during an intervention period from an estimate of the number of crimes that would have occurred had the initiative not existed. The estimate is calculated as the observed number of crimes in the target area in the period prior to the intervention times the ratio of observed crime in the comparison area after and before the intervention. The formula developed by Johnson is:

> *Burglaries prevented = Actual Crime in Action Area (after) – Expected Crime Rate; where the Expected Crime Rate = Crime in Action Area (before) X (Crime in comparison area (after)/crime in Comparison Area (before))*

The above equation was used in the evaluation study of Operation Burglary Countdown to calculate an estimate of the number of residential burglaries prevented in Bentley and Morley and their surrounding suburbs over the 12 months of the pilots. The results are summarised in four three-month quarters in Table 2. Using these figures, Operation Burglary Countdown can be seen to have saved burglaries in Bentley in every quarter for a total of 129 over the period of the program. Unexpectedly, there were also predicted burglaries prevented in the suburbs surrounding Bentley in three of the four quarters for a total of 74 over the full 12 months. In Morley, the impact of the program is less consistent, with small, predicted reductions in burglaries in three of the four quarters but not the final quarter, resulting in a total of -1 over the program period. The evaluation study identified at least two possible reasons for the different results for the two target areas:

- There was a police blitz in Morley just prior to the commencement of Operation Burglary Countdown which reduced the burglary rate to an unusually low number of burglaries (62) in the months used as a comparison for the final quarter of Operation Burglary Countdown, (August – October 2003), as can be seen in Figure 3 (see page 14).
- The government and community engagement in the Local Management Team was less evident and was focused more on long-term strategies in Morley than in Bentley.

Cost–benefit analysis

The financial costs of residential burglary to the community are significant. According to Mayhew (2003), the average cost per reported burglary nationally is estimated at $2400 with the costs being $2000 per incident for residential burglary and $4500 per incident for non-residential burglary. More precise figures were calculated for Bentley and Morley and their surrounds as part of the evaluation study by using actual costs of the police resources in investigating crimes and the cost of property stolen and/or damaged reported to police during the period of the program. For the period of the program, the average was $3900 for Bentley, $2870 for the suburbs surrounding Bentley, $1900 for Morley and $2900 for the suburbs surrounding Morley. These figures are considered to be conservative because they do not include the cost of the justice system or incarceration for apprehended and convicted offenders, increased insurance premiums, the cost of improved security, and the social cost of the intrusion into one's home.

As can be seen in Table 2, it is estimated that in Bentley there were community savings generated of $503 100 and in the surrounding suburbs $212 380 for a total of $715 480. When this is compared to the cost of Operation Burglary Countdown, estimated at $150 000 across the two pilots, there is clearly a considerable net financial benefit to the Bentley community of $428 100. Using Welsh and Farrington's (2000) approach to cost-benefit analysis, Operation Burglary Countdown can be seen to have a benefit-cost ratio ($503 100 ÷ $75 000) of 6.71. In other words, for every dollar invested in Operation Burglary Countdown in the Bentley community a benefit $6.71 was returned. If the diffusion effect in the surrounding area is also included, the benefit-cost ratio increases to 9.54 ($715 480 ÷ $75 000) or a benefit of $9.54 for every dollar invested. In Morley, it appears the program had a positive impact in three of the four quarters but in only one quarter in the surrounding areas. Overall, there was a net negative number of burglaries prevented of -1 in Morley and -137 in the surrounding areas, giving community savings of -$1900 and -$397 300 respectively. This results in a benefit–cost ratio of -0.02 for Morley and -5.32 for the combined area. When the two target suburbs are combined the program achieved an overall savings of $316 280, or a benefit–cost ratio of 2.11. The quite different results in the two suburbs and their surrounding areas indicate that the selection of locations for this type of program is critical.

The evaluation study concluded that in terms of its impact on burglary numbers, the result for Operation Burglary Countdown could not be more dramatic for Bentley; however, it would have been more appropriate to have chosen a different suburb than Morley to test the program. In using a counterfactual analysis, the study provides a useful comparison for the program outcomes and some useful lessons for expansion of the program to other localities.

Beneficial role of the counterfactual in influencing stakeholders

Having explored the theoretical and practical basis for counterfactual analysis, four areas of potential benefit for influencing stakeholders can be identified.

The first is in relation to the impact of counterfactual analysis on how stakeholders perceive the causal relationships that underpin a program's logic model. The research on counterfactual thinking summarised by Mandel, Hilton and Catellani (2005) indicates that an individual's perception of the causes of outcomes is strongly influenced by undertaking counterfactual analysis. The argument is that conducting counterfactual analysis (either quantitative or qualitative) either strengthens or eliminates possible alternative 'counterfactual' causes, thus consolidating the perception of the 'factual' cause. In some ways, this is what a program logic does, although it is rare to see a counterfactual analysis process incorporated in the development of a program logic.

Related to this is the role a counterfactual analysis might play in improving stakeholder understanding of a program's logic model. It would seem to be a useful process for evaluators to work

with stakeholders to identify possible alternative casual relationships. These would initially clarify the causal paths in the chosen logic model as well as provide some direction for the evaluator as to the possibilities of a counterfactual analysis that might be conducted in the evaluation study.

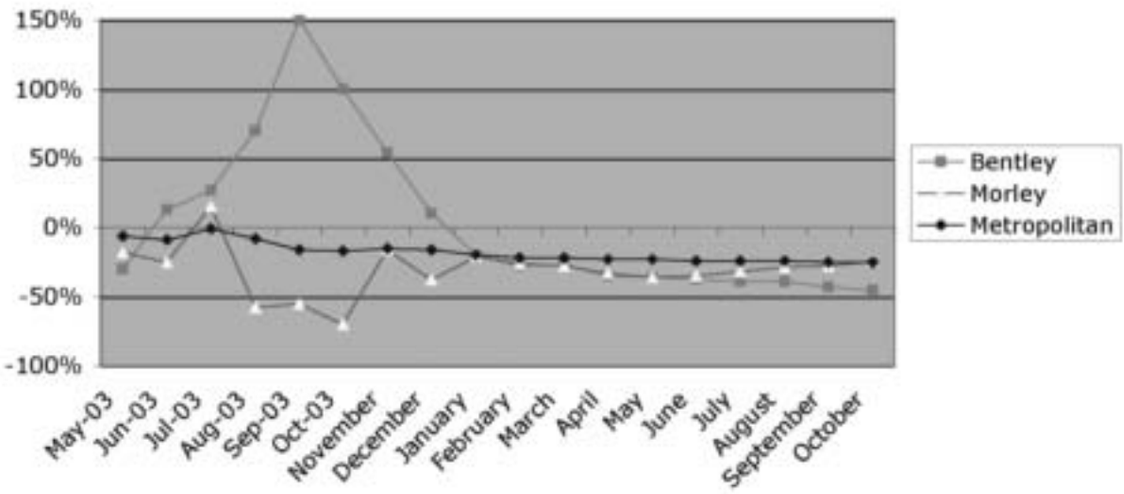
A third benefit of conducting a counterfactual analysis is that it can provide more rigorous and definitive evidence of outcomes. The rigour required in a counterfactual analysis in clearly identifying the outcomes, determining how to measure them accurately, collecting usable qualitative

TABLE 2: BURGLARIES PREVENTED BY OPERATION BURGLARY COUNTDOWN BY QUARTERS

Period	Bentley	Bentley surrounding areas	Morley	Morley surrounding areas	Metropolitan area minus pilot areas
Cost per burglary	*$3900*	*$2870*	*$1900*	*$2900*	
1st quarter (Nov. – Jan.)					
Previous year	163	433	112	404	8476
Program year	130	337	89	405	6786
Change	-33	-96	-23	1	-1690
Burglaries prevented	1	10	1	86	
Costs saved per area	**$3900**	**$28 700**	**$1900**	*-$249 400*	
2nd quarter (Feb. – Apr.)					
Previous year	156	372	100	278	7940
Program year	78	286	54	213	5919
Change	-78	-86	-46	-65	-2021
Burglaries prevented	40	-4	22	-2	
Costs saved per area	**$156 100**	*-$11 480*	**$41 800**	*-$5800*	
3rd quarter (May – July)					
Previous year	156	397	90	347	7403
Program year	79	238	61	199	5282
Change	-77	-159	-29	-148	-2121
Burglaries prevented	34	49	4	52	
Costs saved per area	**$132 600**	**$140 630**	**$7600**	**$150 800**	
4th quarter (Aug. – Oct.)					
Previous year	172	354	62	305	6999
Program year	68	234	72	319	4910
Change	-104	-120	10	14	-2090
Burglaries prevented	55	18	-28	-102	
Costs saved per area	**$214 500**	**$51 660**	*-$53 200*	*-$295 800*	
Program total					
Previous year	647	1556	364	1334	30 818
Program year	355	1095	276	1136	22 897
Change	-292	-461	-88	-198	-7922
Burglaries prevented	**129**	**74**	**-1**	**-137**	
Percentage change in burglary numbers	-45%	-30%	-24%	-15%	-26%
Total per area	$503 100	$212 380	-$1900	-$397 300	
Costs saved overall		$715 480		-$399 200	$316 280

Negative cost values are in italics and indicate no savings.

FIGURE 3: PERCENTAGE CHANGE IN BURGLARY NUMBERS BY TARGET SUBURB AND METROPOLITAN PERTH—MAY 2003 TO OCTOBER 2004



or quantitative data and conducting rigorous and appropriate analysis of the data are further emphasised where the data must also be collected for either alternative program scenarios or non-program areas/individuals.

Finally, a counterfactual analysis can provide very useful evidence for decision-making and very convincing evidence for persuading stakeholders of the outcomes of the program. In Operation Burglary Countdown, the initial results of the evaluation study, including the counterfactual, persuaded the Office of Crime Prevention to expand the program to a third pilot location, Carnarvon, a regional centre in north-western WA. Applying the same counterfactual analysis, but using non-metropolitan WA as the comparison area, revealed very positive results similar to those for Bentley. At the completion of the evaluation study and submission of the final report, the Minister for Justice expanded the program to a whole region in Perth, largely based on the evidence of the evaluation study.

Note

1 These are available online at <http://www.police.wa.gov.au/AboutUs/AboutUs.asp?SearchCrimeStatistics>.

References

Cornish, DB & Clarke, RV (eds) 1986, *The reasoning criminal: rational choices perspectives on offending,* Springer-Verlag, New York.

Cummings, R 2005, *Operation Burglary Countdown: evaluation study final report,* Estill & Associates, Perth.

English, B, Straton, R & Cummings, R 2002, 'Choosing evaluation models for community crime prevention programs', in N Tilley (ed.), *Crime Prevention Studies,* vol. 14, pp. 171–225.

Henderson, M 2002, *Preventing repeat residential burglary,* Commonwealth Attorney-General's Department, Canberra.

Johnson, SD, Bowers, KJ, Jordan, P, Mallender, J, Davidson, N & Hirschfield, AFG 2004, 'Evaluating crime prevention scheme success: estimating 'outcomes' or how many crimes were prevented', *Evaluation,* vol. 10, no. 3, pp. 327–348.

Kvart, I 1986, *A theory of counterfactuals*, Hackett Publishing Company, Indianapolis.

Lebow, RN 2000, 'What's so different about a counterfactual?', *World Politics*, vol. 52, July, pp. 550–585.

Lewis, DK 1973, *Counterfactuals*, Harvard University Press, Cambridge, Massachusetts.

Mayhew, P 2003, 'Counting the costs of crime in Australia', *Trends and issues in crime and criminal justice,* no. 247, Australian Institute of Criminology, Canberra.

Mandel, DR, Hilton, DJ & Catellani, C (eds) 2005, *The psychology of counterfactual thinking*, Routledge, London.

Mohr, LB 1995, *Impact analysis for program evaluation*, Sage Publications, Thousand Oaks, California.

Owen, J 2006, *Program evaluation: forms and approaches,* 3rd edn, Allen & Unwin, Sydney.

Purdon, S, Lessof, C, Woodfield, K & Bryson, C 2001, *Research methods for policy evaluation*, Department for Work and Pensions, Research Working Paper no. 2.

Rosenbaum, DP 2002, 'Evaluating multi-agency anti-crime partnerships: theory, design and measurement issues', in N Tilley (ed.), *Crime Prevention Studies,* vol. 14, pp. 171–225.

Stufflebeam, DL 2001, *Evaluation models, New directions in program evaluation*, no. 89, Jossey-Bass, San Francisco.

UK Prime Minister's Office 2004, *Strategy survival guide: gathering evidence,* viewed 24 July 2006, <http://www.strategy.gov.uk/downloads/survivalguide/skills/eb_counter.htm>.

Welsh, CA & Farrington, DP 2000, 'Monetary costs and benefits of crime prevention programmes', in M Tonry (ed.) *Crime and Justice: A review of research*, vol. 27, pp. 305–361.

Worthen, BR, Sanders, JR & Fitzpatrick, JL 1997, *Program evaluation: alternative approaches and practical guidelines*, 2nd edn, Longman, White Plains, New York.