EVALUATION OF A MENTORING PROGRAM FOR YOUNG JUVENILE CRIME OFFENDERS – CHALLENGES FOR EVALUATION DESIGN

Chris Milne, ARTD Consultants, and
Bruce Flaherty, Crime Prevention Division, NSW Attorney General's Department.

Programs to prevent juvenile crime offer a number of challenges to government, communities, service agencies and criminologists. There are also plenty of challenges in evaluating crime prevention programs, as outlined in this paper on the design of an evaluation for a pilot mentoring program. ARTD was commissioned to evaluate the program at around the time the program was being established, and since then have been working on the evaluation design in conjunction with the Crime Prevention Division (CPD) of the NSW Attorney General's Department and the program staff. This paper briefly describes the program and the purposes of the evaluation. It then discusses the design of the evaluation, and in particular challenges about effectiveness and replication, as a contribution to the debate around the evaluation of crime prevention.

A MENTORING PROGRAM FOR CRIME PREVENTION

Concern with high rates of juvenile crime has led to interest in innovative approaches to crime prevention in Australia (National Crime Prevention, 1999). One approach is "*mentoring*", where a young person meets regularly with an adult in order to produce a stable and continuous friendship. This relationship is assumed to reduce anti-social behaviours including offending. This type of mentoring has become a significant intervention with at risk" young people in the USA and the UK (Grossman and Garry 1997, Audit Commission 1996, 1998) For example, the U.S. Office of Juvenile Justice and Delinquency Prevention has funded over 120 projects for around 3000 young people through its Juvenile Mentoring Program (JUMP) between 1994 and 1998 (US Department of Justice 1998)

In NSW, the Department of Juvenile Justice and the Crime Prevention Division is trialing a mentoring program as a crime prevention initiative, with a three year pilot program being developed in two areas: Parramatta in Western Sydney and the Coffs Harbour/Clarence Valley region on the mid north coast of NSW. Young people who have recently committed a criminal offence will be offered a mentor who can give them positive adult guidance, support and friendship. The program will be targeted particularly to those young people where there is an absence of such relationships in their lives. The program objectives are to ensure that through the provision of appropriate mentoring, young people at risk will be diverted from further involvement with the juvenile justice system and will become more integrated into the community. The program is expected to cater for a total of about 200 young people, or 35 per year in each of two locations.

The program has the support of all the relevant agencies, with a Steering Committee including representatives of the Department of Juvenile Justice, the Youth Justice Conferencing Directorate, the Department of Community Services, the NSW Police Service, and the Office of Children and Young People within the Cabinet Office.

With the program aiming to target recent offenders, a ready-made target group was provided by the NSW Young Offenders Act (1997), which established a system of official warnings, cautions and youth justice conferences as alternatives to court appearances for young people. The Steering Committee determined that the target group for the mentoring project would be young people who were cautioned by Police or referred to a Youth Justice Conference. The Committee also decided that

participation was to be voluntary and should be seen as a positive opportunity rather than a punishment. Mentoring was not to be a condition of a caution, nor part of the outcome plan from conferencing (Crowley and McInytre, 1999).

The YWCA of Sydney, a charitable non-government organisation with a range of programs for children, adolescents and women, was selected to develop and conduct the project. The YWCA had been running its Big Sister Big Brother (BSBB) mentoring program in Sydney over the last 20 years. BSBB broadly follows the Big Brother Big Sister model, a program that is widely established and well regarded in the USA, with over 500 agencies and a reported 75,000 active matches in 1995 (Tierney and Grossman, U.S. Department of Justice, 1998).

The mentoring program for young offenders, now named One$_2$One, is currently being established by the YWCA. While the YWCA's BSBB has worked with many young people at risk, One$_2$One will have some important differences: young people will be referred from the justice system, and they will be offenders and probably older than the current BSBB group. Some features of One$_2$One will be able to be readily transferred from BSBB (eg selecting and supporting mentors, making matches). Other features will need to be developed or adapted as the program evolves (eg referrals from Police and conferences, recruiting mentors to work with young offenders).

An initial scan of the field by ARTD found that while mentoring programs are widely used with young people at risk, they can vary in many key features (eg Benioff, 1997). The specific features of the One$_2$One model are shown in Table 1.

Table 1 Features of the One$_2$One model for mentoring

Rationale	Reduce the risk of a further drift into crime by providing young people with a positive adult role model/mentor who provides consistent friendship (Crowley and McIntyre, 1999).
Goal	Produce "performing matches" between young people and volunteer mentors
Role of mentor	To provide young person with a structured and supported one-to-one friendship with a caring adult over a period of at least twelve months.
Staffing	Professional project officers to manage recruitment, referrals, matches, mentor training and support, community links. Part time family support worker to work with families.
Target group (young people) and referrals.	Young people who have been charged with an offence and cautioned by Police or referred to Juvenile Justice Conference, who exhibit risk factors in relation to offending, and who are seen to have the capacity to benefit from mentoring. Referrals are made by police or conference administrators.
Recruitment and screening of mentors	Volunteers are sought from the local community. Screening and training takes four months using BSBB resources and standards. Around 10% of initial inquiries are expected to result in active mentors.
The match	The young person is assessed and then matched with a mentor who has suitable interests and capacities. Young person, parent and mentor sign match agreement. The mentor and the young person spend 2 to 5 hours per week, usually at the weekend, doing activities together. The match is for a minimum period of twelve months.
Ongoing support of match	The project officer reviews progress of the match through assessments at 3 months and 6 months. Project officer also supports the mentors through weekly contact in first 3 months. Mentors are expected to attend volunteer meetings each month.
Other aspects	X The young person's family is assisted by the family support worker in order to support the young person's continuing participation X Group activities are conducted which the mentor and young person are encouraged to attend. Young people are encouraged to feel that they belong to the One$_2$One program.

In thinking about the evaluation of One$_2$One, and particularly what is understood by effectiveness, it is important to be clear on the underlying logic of the mentoring strategy. Through discussions with the One$_2$One staff, we have developed a summary statement of the logic (Figure 1) and an outcomes hierarchy (Figure 2):

Figure 1: The logic of One$_2$One

For certain groups of young offenders who have been cautioned or conferenced, the experience of a positive and ongoing relationship (a "performing match" based on weekly meetings over twelve months) with an appropriate adult (a volunteer who has been carefully selected, trained and supported) will reduce risk factors related to offending and ultimately reduce their likelihood of re-offending. The young person's involvement in the mentoring relationship is more likely to be successful if the family of the young person is assisted to support the young person's involvement; and if the young person and mentor participate in organised group activities with other pairs.

In addition to the mentoring strategy itself, One$_2$One also has strategies to establish the program within the selected communities, and to promote appropriate referrals from police and conferences. Both these elements are relevant to questions about replicating the program in other communities.

CHALLENGES FOR THE EVALUATION

The Steering Committee had two main purposes in evaluating the program. Firstly, as a pilot program, evaluation was needed to inform decisions about the future use of mentoring for crime prevention. A *summative* evaluation was required to assess the overall effectiveness of the program in achieving its outcomes, and more specifically, to examine how well this model of mentoring worked for different types of young people, and in different social contexts. Further, as this was a new and evolving program, the evaluation had to describe and assess how well a mentoring program could be established and implemented with this target group, and in the selected communities.

Secondly, a *formative* evaluation process was needed to provide information to the program team and Steering Committee during the course of implementation about where processes could be improved. Obviously there is overlap between these two elements of the evaluation, and an important ingredient for both is a monitoring system which could provide useful data on program implementation and participation and outcomes for young people.

One$_2$One is a new and evolving program with a range of stakeholders with differing expectations about outcomes. CPD and the Dept. of Juvenile Justice are specifically looking for its contribution to reduced offending. On the other hand, BSBB see the achievement of a performing match as the critical outcome. Both stakeholders are interested in whether One$_2$One can be implemented with young offenders - will they agree to participate and continue to be involved? These outcomes are not unrelated, and the links were documented in the outcomes hierarchy (Figure 1).

Given its purposes, the challenges in designing an evaluation strategy that we have been dealing with are:

X   *How to assess the effectiveness of the program **as a crime prevention strategy;***

X   *How to assess the effectiveness of the program **in reducing risk factors** for the participants;*

X   *What information is required for decisions to be made about **replication**;*

X   *What are the most appropriate **data collection methods***

Before looking at each of these, we should state our general perspective on evaluation, which since the work of Pawson and Tilley (1997) we find we can describe as favouring scientific realism, or "realistic evaluation" rather than experimental or constructivist perspectives. We are interested not only in the outcomes that can be attributed to an intervention, but also in better understanding how the intervention works, with whom, and in what contexts, or in the terminology of realistic evaluation, in specific configurations of contexts, mechanisms and outcomes. It has also been reassuring to find that this perspective has been broadly regarded as the most appropriate for the evaluation of early interventions into crime prevention in Australia (National Crime Prevention, 1999, p.97-99).

*Challenge One - How to assess the effectiveness of the program **as a crime prevention strategy**;*

The government agencies want the evaluation to specifically assess the program's contribution to reduced offending. The easier task will be to collect data on any further offending over the period of their mentoring. The harder task will be to interpret the data and link it to the intervention within the period of the evaluation, in the light of a host of countervailing issues such as - many young people are not likely to re-offend anyway; young people's offending behaviour tends to change over time as they grow older; the impacts may be longer term and show up outside the period of the evaluation; mentoring may be just one of a number of interventions over the period, offending may be affected by changing local conditions such as policing and new opportunities; and so on.

There may be some scope for comparative data on offending. Our initial analysis suggests that for every young offender who is cautioned or conferenced at the two sites, perhaps 20-30% will meet the selection criteria for referral to One$_2$One, but possibly only 5% will be able to be matched with a mentor. The police data system keeps records on any further offending, so it may be feasible to compare the pattern of offending for the mentored group with those who met the selection criteria but who were not able to be mentored, for example, because of the unavailability of a mentor.

We are also addressing this challenge in other ways. Firstly, there is strong evidence that offending is associated with a number of risk and protective factors (National Crime Prevention, 1999), and the JUMP programs in the USA have generally had this focus (U.S. Department of Justice, 1998). One$_2$One aims to reduce some of these risk factors (such as low self esteem, poor social skills, alienation, poor attachment to school) and provide a protective factor (supportive relationship with adult outside the family). If One$_2$One can be shown to successfully contribute to these outcomes for certain groups of young people in certain circumstances, then the evaluation will be able to infer that it will reduce the likelihood of their offending, though not in any way that could be quantified. At the same time, assessing the impact of mentoring on the risk factors has its own challenges, which we discuss further below.

Secondly, it can be argued that One$_2$One is directly linked with crime prevention because of its target group and referral process. One$_2$One is not a general support program for young people, but specifically targets young people who have already taken some steps down the path of offending, so it is already in the domain of crime prevention. For the young people, participation in the program should be firmly linked with their offending, with referral made by the Police or conference administrators. From this perspective, the question about effectiveness becomes *How well is One$_2$One is able to gain and*

maintain the participation of this high priority target group. Thus the evaluation needs to document the pattern of participation in the program and make comparisons between groups of participants (eg age, gender, ethnicity, type of offence) as well as between participants and all young people who are cautioned or conferenced at the two locations.

Another view of the effect of One$_2$One on offending behaviour will come from the perceptions of the young people, their parents and other key people with whom they are involved. Interestingly, we have already found much enthusiasm for mentoring in interviews with Police and others working in juvenile justice. For those at the coal face, mentoring has strong face validity. We plan to collect these perceptions and use them in case studies of young offenders which describe their "stories" about mentoring and their involvement in offending.

Challenge Two - How to assess the effectiveness of the program **in reducing risk factors** *for the participants;*

Collecting information on changes to risk factors offers some methodological challenges. Young people will come to the program with a diverse array of backgrounds and circumstances. For One$_2$One, the mentoring relationship is a friendship that can have many possible positive impacts, but does not have specific objectives – it is not case management where certain outcomes or objectives are defined in advance. To address this, our approach will be to compare changes in the five or six main risk factors for all young people when they first enter the program and when they have been in it 9-12 months, as well as open ended recording of other risk factors and issues.

However, the design challenge is a broader one - how to attribute any changes in the risk factors to One$_2$One, since some changes in these outcomes are likely to occur anyway due to young people growing older, changes in life circumstances, and other interventions. After all, mentoring may account for 5% of a young person's waking hours, compared with, say, school at 30%.

We explored the option of using comparison groups to increase confidence in attributing any changes in risk factors to the program. The one study we have found which did this examined eight established Big Brother Big Sister mentoring programs for at-risk young people in the USA. 959 young people who applied were randomly allocated to a treatment group, for which BBBS matches were attempted, or to waiting lists, and both groups were assessed by the BBBS projects. A follow up assessment was made after 18 months. The findings demonstrated some benefits of mentoring. For example, the young people with mentors were less likely to report initiating drug use (46%) or alcohol use (27%), hitting someone (32%), or to skip a day from school (52%) (Tierney and Grossman 1995). The study is important in that it supports the program logic of One$_2$One by demonstrating that mentoring using the BBBS model can impact on risk factors.

However, such an approach does not seem feasible for One$_2$One. Waiting lists are not anticipated, indeed an important feature is to introduce mentoring soon after the young person's involvement with the criminal justice system. We considered using other comparison groups, such as the young people who met the selection criteria but were not matched, which we discussed above in relation to re-offending rates. However, the detailed assessment and follow up of such a group is well beyond the resources and scope of the evaluation (see Challenge Four). Another issue is that of program fidelity - the One$_2$One "treatment" may not be standard, as elements of the intervention may change over the period as more is learnt about its strengths and weaknesses.

In considering risk factors, the evaluation focus for One$_2$One as a pilot program is more to explore how mentoring can be implemented with a different population - young offenders in NSW. We plan to examine its effectiveness in addressing risk factors by (a) developing a program logic for the program which shows different levels of outcomes, factors and contexts, and collecting data on each of these to see how far up the hierarchy One$_2$One can be effective, and (b) undertaking case studies which explore the effectiveness of mentoring on individual participants and their families.

Challenge Three - Determining the information required for decisions to be made about **replication**.

Programs are piloted for different purposes, and for One$_2$One it is primarily to explore the unchartered waters of mentoring with young offenders in NSW. Exploratory pilots raise special issues for evaluation, since the outcomes from the intervention - in this case, mentoring - are often produced from diverse patterns of implementation, and in particular communities and circumstances (Brooks and Milne 1997). The replication question is *If the mentoring was implemented in a different way, or within a different community, or with a different group of young people, would it produce the same outcomes?* While there will always be uncertainty in answering this question, Tilley (1996) proposed a conceptual framework to reduce it by documenting key features of the program. The framework is based on the mechanisms that were used, the contexts that triggered them, and the outcomes that were produced.

For the evaluation, this approach to the replication issue requires the collection of systematic data on the specific interventions (mechanisms), the conditions that were associated with producing the outcomes (contexts) and the outcomes at different levels. As a tool to identify the data that is needed, we will use the outcome hierarchy (Figure 2) to develop a detailed program logic matrix covering outcomes measures, mechanisms and contexts at each level.

Challenge Four - Identifying the most appropriate **data collection methods**

Decisions about data collection methods require balances between depth, spread, effort, budget, sustainability and reliability, whilst taking care to limit bias, side effects and potential harm to participants. For this evaluation, a crucial issue is to minimise any intrusion into the relationships between young person, mentor and One$_2$One. At the same time, trust and rapport is needed with these young people before they are open to questions about personal matters.
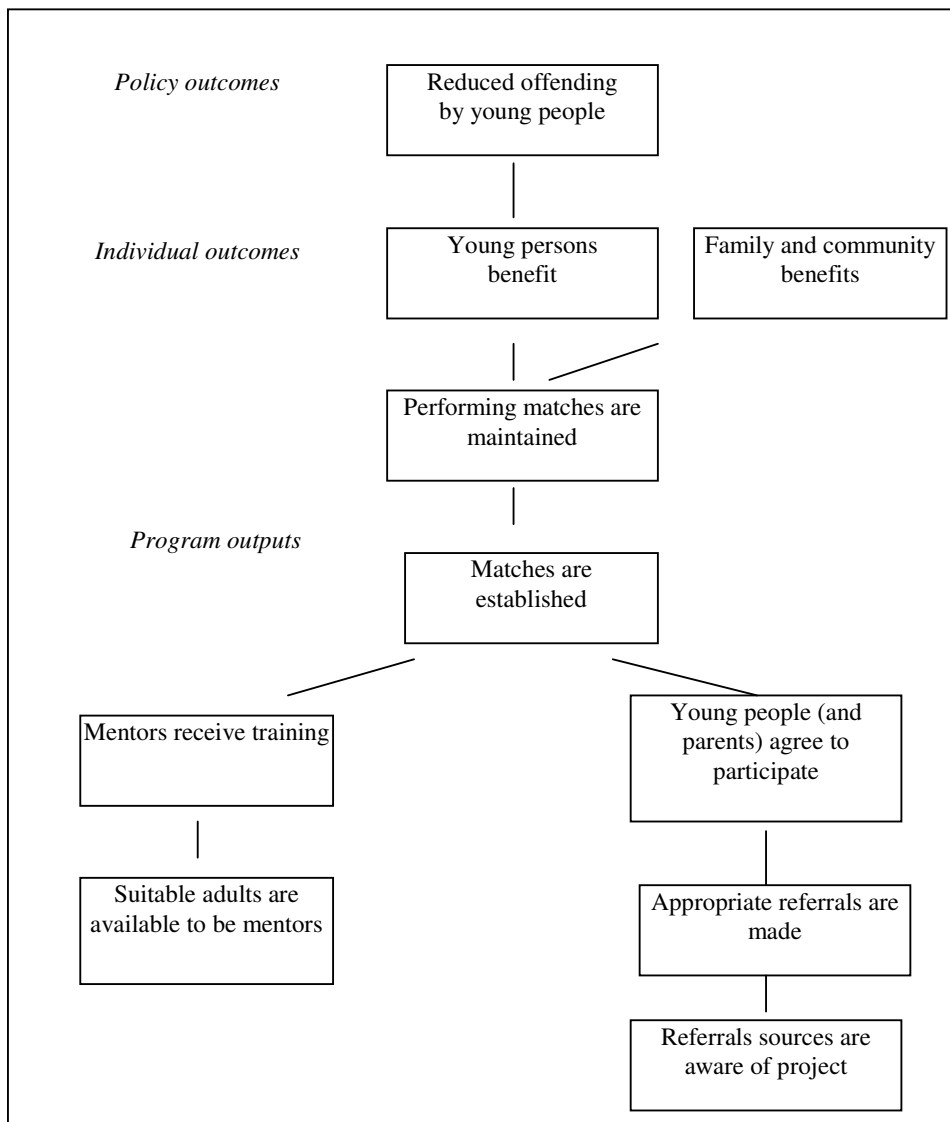
The formative evaluation component gives us the scope to work together with One$_2$One to design these measures and to utilise the program processes as much as is feasible for data collection. Our aim is to ensure that the data collected will be useful for One$_2$One and the evaluation, cost effective and sustainable. We have agreed that One$_2$One will collect monitoring data on referrals, participation and participant characteristics. As part of their program assessment and review responsibilities, the One$_2$One project officers will make standardised assessments of risk factors.

This data will be validated by other measures. ARTD will interview all the young people and mentors, but focus on their satisfaction with the mentoring processes, and also collect corroborative data from schools and police. In addition, a purposive sample of 20 case studies will be undertaken by ARTD, involving in-depth interviews with young people, their families and other stakeholders. We feel this multi-layered approach to data collection will be consistent with the One$_2$One program yet be sufficiently triangulated and comprehensive to provide the data needed to address the three other challenges.

CONCLUSION

At this stage we are reasonably confident that the evaluation design can address the challenges we have identified, particualry in assessing the effectiveness of One$_2$One for crime prevention and for reducing risk factors, and in providing information for decisions about replication. What may be more interesting for the evaluation as One$_2$One rolls out over the next three years is how well the design will be able to address the challenges we have not anticipated.

Figure 2: **Outcomes hierarchy for One2One Program**

Policy outcomes — Reduced offending by young people

Individual outcomes — Young persons benefit | Family and community benefits

Performing matches are maintained

Program outputs — Matches are established

Mentors receive training | Young people (and parents) agree to participate

Suitable adults are available to be mentors | Appropriate referrals are made

Referrals sources are aware of project

References

Audit Commission (1996) Misspent Youth - Young People and Crime, National Report, Audit Commission, London, UK.

Audit Commission (1998) Misspent Youth '98 The Challenge for Youth Justice, Update, June 1998, Audit Commission, London, UK.

Benioff, Sarah (1997) A Second Chance - Developing Mentoring and Education Projects for Young People, Commission for Racial Equality and Crime Concern, London, UK.

Brooks, M. and C. Milne (1997) *It'll be alright on the night: Issues in the design and evaluation of pilot programs*, Proceedings of the Australasian Evaluation Society International Conference, October 1997, Adelaide, Australia.

Crowley, Desmond and Shauna McIntyre (1999) One$_2$One - A Mentoring Project for Young Offenders Paper presented to Australian Institute of Criminology Conference, "Children and Crime: Victims and Offenders", 17-18 June, 1999, Brisbane, Australia.

Grossman, Baldwin Jean and Eileen M. Garry (1997) Mentoring - A Proven Delinquency Prevention Strategy. U.S. Department of Justice, Office of Justice Programs, Office of Juvenile Justice and Delinquency Prevention, OJJDP Bulletin April 1997, Washington, USA.

National Crime Prevention (1999) Pathways to Prevention: Developmental and Early Intervention Approaches to Crime in Australia. National Crime Prevention, Attorney-General's Department: Canberra, Australia.

Pawson Ray and Nick Tilley (1997) Realistic Evaluation, Sage, London, UK.

Tierney, J. P., and Jean Grossman (1995) Making a Difference: An Impact Study of Big Brothers/ Big Sisters. Public/Private Ventures, Philadelphia, PA, USA..

Tilley, N. (1996) "Demonstration, Exemplification, Duplication and Replication in Evaluation Research", Evaluation, Vol 2 (1) pages 35-50.

U.S. Department of Justice (1998) Report to Congress - Juvenile Mentoring Program (JUMP) U.S. Department of Justice, Office of Justice Programs, Office of Juvenile Justice and Delinquency Prevention, Washington, USA.