

On being realistic about reducing the prevalence and impacts of youth sexual violence and abuse in two Australian Indigenous communities

Nick Tilley University College, London n.tilley@ucl.ac.uk		
Susan Rayment-McHugh Griffith University	Stephen Smallbone Griffith University	Martina Wardell Department of the Prime Minister and Cabinet
Dimity Smith Griffith University	Troy Allard Griffith University	Richard Wortley University College, London
Donald Findlater Lucy Faithfull Foundation	Anna Stewart Griffith University	Ross Homel Griffith University

Keywords: realist evaluation, realistic evaluation, realist design, youth sexual violence and abuse, community safety evaluation, evaluation in Australian Aboriginal and Torres Strait Islander communities, Griffith Youth Forensic Service Neighbourhoods Project

Abstract

Social interventions, like medical ones, can produce negative as well as positive outcomes. It is important for policy and practice to learn what works, what doesn't work, and what produces unintended effects, for whom and in what contexts. This is the task of realist evaluation. The formulation and evaluation of programs aiming to deal with problems in Australian Aboriginal and Torres Strait Islander communities face a number of practical, conceptual and methodological problems. Here, realist methods for the design and evaluation of promising programs from which transferable lessons can be derived are discussed in the context of an initiative aiming to reduce the prevalence and impacts of youth sexual violence and abuse. Tentative conclusions are drawn for what this might mean for programs targeting similar problems elsewhere.

Introduction and background

Griffith Youth Forensic Service Neighbourhoods Project (GYFS-NP) aims to reduce the prevalence and impacts of youth sexual violence and abuse (YSVA) in two Australian communities – a remote Aboriginal community, and a culturally diverse suburban precinct within a regional city. It aims to do so by engaging closely with the local communities, measuring various aspects of the problem in a variety of ways, and developing, implementing and evaluating a suite of locally-tailored interventions. Because of the context-specific nature of the targeted problems, these particular interventions may not be directly transferrable to other sites or to other related social problems. At the same time, the project aims to develop and test an over-arching prevention model that is transferable to a wide variety of places, problems, and contexts.

YSVA has emerged as a hitherto only partially recognised problem in some Aboriginal and Torres Strait Islander (Indigenous)¹ communities. It is experienced in remote communities as well as within communities embedded in towns and cities. It takes the form of rape, prostitution, ‘rough sex’ in which girls appear to be resigned to being treated as objects of sexual satisfaction to boys, inculcation of children into highly sexualised peer groups, sexual teasing, self-abuse, and inappropriate touching.

The project whose evaluation is the concern of this paper began with an exploration of a range of data on the extent of YSVA in the project’s target areas, and particularly amongst the Indigenous people residing in these areas. Data were assembled going back for as long as a decade. Rates of pregnancy amongst young girls, sexually transmitted infections, and recorded sex offences were many times higher within the project’s target areas and amongst Indigenous residents than amongst the general population. Systematic observations were also made at key public locations within the project areas where it was thought that antisocial behaviour might be taking place, when it was deemed safe enough, not in the expectation that YSVA would be seen directly but to gauge the number and basic attributes of those present, what they were doing, and the presence or otherwise of formal and informal guardians. The project is led by psychologists based at Griffith University whose work has specialised in the assessment and treatment of court-referred youth sexual offenders. Work with these youth strongly suggested that the incidents coming to the attention of the authorities, particularly in the two target communities, represent only a small fraction of the total number of incidents occurring. Indeed it suggested that some forms of YSVA might be endemic in the two locations of concern (Smallbone & Rayment-McHugh, 2013).

Following an earlier study that aimed to empirically examine the scope, dimensions and dynamics of YSVA in these two communities (Smallbone, Rayment-McHugh, & Smith, 2013a), Australian Government funding was secured to support a three-year program that would devise, deliver and evaluate strategies to reduce the extent and

1. We are mindful of sensitivities about terms used to refer to Australian Aboriginal and Torres Strait Islander peoples. We have used the term ‘Indigenous’ in this paper because it accords with references to policies and other documents.

impacts of the problem. It was agreed that a realist approach would be taken to the evaluation. Project funding provides for two full-time clinicians, a full-time researcher, an international advisory group, travel and associated expenses. It does not provide for the costs of most of the interventions that are likely to be formulated during the project. Though some of the interventions will be provided directly by the two full-time clinicians, others will depend on third party agreement and action.

Following public health principles, the project adopts a broadly preventative approach. This includes primary prevention – measures aimed at the general population and background conditions that enable or foster YSVA; secondary prevention – measures targeting the types of setting that are known to foster YSVA and the groups known to be at high risk of YSVA offending or victimisation; and tertiary prevention – measures targeting existing sites for YSVA and known perpetrators or victims, to try to prevent further incidents. Prior work of the clinicians leading the project had focused on the tertiary prevention of continued sexual violence and abuse by known youth offenders (Smallbone & Rayment-McHugh, 2013; Smallbone, Rayment-McHugh, Crissman, & Shumack, 2008; Smallbone, Rayment-McHugh, & Smith, 2013b). However it soon became clear that even if these clinical interventions successfully prevented further offending among individual referred youth, this would have a very limited overall effect on the prevalence and impacts of YSVA in these communities. Serious concerns remained that other young people, exposed to the same situational, family, peer, organisational and neighbourhood risk factors, would go on to offend or be victimised unless their exposure to these risk factors was reduced or prevented.

A ‘Haddon matrix’ is being used to capture the different types of prevention and targets of intervention. Table 1 shows the matrix and provides some illustrative examples of the types of measures that might be considered (see Smallbone, Marshall, & Wortley, 2008, for a fuller discussion of the matrix).

Table 1: Haddon matrix for YSVA with examples of measures that might be considered

	<i>Primary Prevention</i>	<i>Secondary Prevention</i>	<i>Tertiary Prevention</i>
Offenders/ Potential Offenders	<ul style="list-style-type: none"> • Reduce exposure to known developmental risk factors • Introduce school-based sexual ethics programs 	<ul style="list-style-type: none"> • Re-engage school-disengaged youth • Therapeutic services, particularly for boys exposed to known risk factors 	<ul style="list-style-type: none"> • Incapacitate most prolific/ serious/ influential youth offenders • Expand offender rehabilitation services (YSVA specific + serious delinquency)
Victims/ Potential Victims	<ul style="list-style-type: none"> • Reduce exposure to known developmental risk factors • School-based resilience-building programs 	<ul style="list-style-type: none"> • Interventions with at-risk girls (personal safety, guardianship, sex education) 	<ul style="list-style-type: none"> • Improve reach and effectiveness of victim support and treatment services • ‘Cocoon’ the most vulnerable victims
Situations	<ul style="list-style-type: none"> • Create safe, attractive places for children and youth • Increase legitimate use of public spaces 	<ul style="list-style-type: none"> • Improve natural surveillance in ‘at risk’ places • Increase planned/ legitimate/ supervised activities in ‘at risk’ 	<ul style="list-style-type: none"> • Improve targeting of police patrols (hot spots; hot times) • Disrupt problem youth group activities/ movements
Community	<ul style="list-style-type: none"> • Mobilise and focus community concerns about YSVA • Parenting programs tailored for the local context 	<ul style="list-style-type: none"> • Responsible bystander training (youth and adults) • Problem-solving with community leaders to reduce barriers to community guardianship 	<ul style="list-style-type: none"> • Mobilise and focus community concerns about YSVA • Community engagement focused on improving extended guardianship

The broad methodology for the development of GYFS-NP has been one of problem-solving. This involves breaking problems down into their constituent parts and then working out what might most plausibly be put in place to deal with them. The process is sometimes described in SARA terms (Eck & Spelman, 1987). The 'S' refers to 'scanning', which comprises a broad effort to assemble evidence on the nature and extent of the problem. The pre-project work involved scanning. The first 'A' refers to 'analysis', where hypotheses relating to the conditions giving rise to the problem are formulated and tested. The 'R' refers to 'response', where strategies and tactics to address the problem based on the analysis are devised and plans made for their implementation. The responses will fit into the Haddon matrix. The second 'A' refers to 'assessment', where the processes involved in practice and the outcomes ultimately achieved are identified by systematic research and evaluation so that transferable lessons can be learned. In practice, although the S→A→R→A sequence looks linear the processes are iterative where work at later stages lead to examination and reworking of earlier ones (Sidebottom & Tilley, 2011).

Local Implementation Groups (LIGs) comprising local community representatives and key local service providers have been established in the project communities to advise on the appropriateness and implementation of the suggested responses that emerge from the scanning and analysis. Negotiations are also being undertaken with third party partner agencies and organisations, whose participation will be necessary to deliver many of the measures that are likely to be suggested.

As this paper is being written scanning has been undertaken and analysis with a view to the formulation of responses is underway. The first interventions have been agreed and are in the early stages of implementation. No evaluation has yet been undertaken. Our focus in the present article is on the planned provision for realistic and realist evaluation. We set the scene with some brief comments about the expectations of government agencies concerning the evaluation of funded Indigenous community safety projects. We then set out a number of challenges to the evaluation of crime prevention projects generally, and to the present project specifically. In the main section of the paper we work through an example of how we are going about the task of applying the realist method to specific sub-problems associated with youth sexual violence and abuse.

Government interest in evaluation: Why evaluate?

'Value for Money' is a key principle that underpins all decision-making regarding the expenditure of public money. In Australia, the Commonwealth Grant Guidelines set out a number of key principles for grants administration, including 'Outcomes Orientation'. Put simply, the Guidelines articulate that agency staff should determine what change is expected as a result of a granting activity and then measure the actual outcome.

A strong focus on evaluation in the Indigenous community safety sector is deemed imperative not just in terms of accountability for public expenditure. It is also impelled by the severity of the issues to be addressed and the paucity of existing evidence about the outcomes of related initiatives. Indigenous people are consistently over-

represented across all Australian criminal justice systems, both as victims and offenders, with direct ramifications for individuals, families and communities (Allard, 2010, 2011; Weatherburn, 2014). This over-representation is also a barrier to improving outcomes in a range of other areas including health, education and employment. The situation is compounded by significant gaps in the evidence base to guide policy and programs. The “*What Works to Overcome Indigenous Disadvantage – Key Learnings and Gaps in the Evidence*” report produced by the Closing the Gap Clearinghouse (2013) identified a large number of studies relevant to community safety. However, it also noted that the vast majority of these studies were descriptive only, with just two projects demonstrating positive outcomes.

Funding was allocated to GYFS-NP primarily to address a significant problem that has historically been insufficiently recognised, acknowledged or addressed, namely youth sexual violence and abuse. This problem undoubtedly exists in all communities, but seems to be especially prevalent in some Indigenous communities. Whilst value for money clearly matters, the issue is also important in its own right.

Community safety evaluation problems and how they apply to GYFS-NP

While knowing the overall value for money from different interventions is an admirable objective, it is important to be realistic about what is possible. The evaluation of criminal justice interventions faces major challenges in all settings. These problems are especially acute when some crime types, such as YSVA, are being addressed. In Indigenous community settings there are also some special difficulties.

This section sets out thirteen major challenges for evaluating crime prevention projects in general, and those focusing on YSVA in particular. At first glance these challenges may appear to present insurmountable barriers to the task evaluating outcomes of such a project. In the later sections of the paper we hope to show that progress, albeit sometimes slow and unsteady, can be made through a systematic process of theorising and empirical testing.

1. *Rare events.* Some crimes, including some forms of YSVA in Indigenous communities (e.g. stranger rapes), are high impact but rare events (Laycock, 2013; Tilley, 2009a). Obtaining meaningful numbers for statistical analysis aimed at measuring short-term outcome effects may be unrealistic.

2. *Multiple interventions.* High crime neighbourhoods typically face multiple problems in education, health and social welfare. For this reason they are often sites for many social programs. A program aiming to address one problem is liable to have an effect on others also. Teasing out the independent net effects of individual initiatives, where many are operating simultaneously and where new ones are being added and old ones fold, is often not possible using standard before and after, experimental and quasi-experimental methods. The YSVA initiative focused on here is a case in point. Both GYFS-NP sites have multiple on-going and new initiatives focused on health and social welfare as well as initiatives focused on other crime problems.

3. *Multiple components in single programs* (Tilley, 1996). Individual crime prevention programs tend to involve a range of components. They often include one or more interventions aimed at awareness-raising, opportunity-reduction, lessening crime motivation, and community capacity building. These composite measures are typically put in place by different agencies. Working out which of the measures are producing positive, negative and nil effects is highly challenging. Net impacts, if they can be assessed at all, may mask varying types of impact in different sub-groups. GYFS-NP will include multiple interventions.

4. *Dodgy and confidential data*. Counting crimes is always difficult. Not all crimes are reported and of those reported not all are recorded (Tilley, 1995). Moreover, rates of crime reporting and recording change. The reasons for change may or may not be related to the content of the program. Raising awareness of YSVA amongst the community and amongst those to whom it is reported is liable to affect the rate at which it is reported and recorded (Farrell & Buckley, 1999). This has largely been the experience of efforts to take domestic violence more seriously, for example. There are alternatives to recorded crime, but these present their own problems. Victimization surveys, which are an obvious possible choice, are expensive, technically difficult to mount, risk responses that are not always candid, and have to be very large to capture relatively rare events. Where administrative records may be usable, for example from police, health or education, for data protection reasons they are often difficult to access at a level of detail which make them usable for evaluation purposes. In the case of YSVA this is an acute problem, not only for evaluation but also for analysis, where the details relating to subsets of young people are crucial.

5. *Adaptation*. Human beings are active agents who behave intentionally and knowingly in response to their situations and the resources they have to hand. They are liable to adapt to a change in response to new resources offered by a program or new understandings prompted by it. This, alongside multiple interventions and multiple components of single interventions, brings a characteristic complexity to the ways in which interventions produce changing and often diverse outcomes (Pawson, 2013; Byrne & Callaghan, 2013). An 'arms race' has been described for some situational measures, for example, whereby over time the offender community innovates in response to measures introduced to make offending more difficult or risky (Ekblom, 1997). Likewise those attempting to reduce crime will then try to find further means of making the crime more risky or less rewarding. With YSVA it is possible to envisage measures that will make it more difficult to find an opportunity for a sexual assault in a crime hotspot, but over time some offenders may find a way round them. The short and long term effects are liable to change so that what was effective at Time One may no longer be effective at Time Two.

6. *Rubbernecking*. Agencies in crime prevention, as in other fields, continually and commendably look for ways of improving their services to the community (Tilley, 2004). There is often also a strong oral culture where new and promising ideas spread rapidly (even poor ones can spread quickly if they have sufficient surface plausibility). This means that an initiative designed for implementation in one area is apt to be appropriated fully or in part in others also. This is liable to undermine comparisons

between experimental and control areas. Even where similar initiatives are not applied some compensatory alternative may be introduced for ‘controls’, again undermining the validity of any control/experimental group comparison. It is unclear yet whether or how this might apply to GYFS-NP.

7. *Unique conditions.* It is trite but true that each individual and each context is unique. Moreover each individual and each context undergoes continual change. Thus, what is in place for one individual at one moment is not precisely the same at the next (Cartwright & Hardie, 2012). Yet what is of interest in the evaluation of programs is not primarily what worked in the past at a specific time for specific groups but what will work now and in the future for the groups whose problems are being addressed. If community safety evaluations are to be useful, they have to be pitched at a useful level, one that transcends the specific. But the evaluation has also to avoid the banalities of the very general, for example that reduction in opportunity can reduce crime. Even if GYFS-NP has a net effect amongst those to whom it is applied, unless there is a reason to believe that similar effects can be achieved elsewhere little of use will have been learned. The inductivist’s error is to assume that what went in the past will go just the same in the future. Evaluations need to achieve an appropriate middle-range level of generality. Few do. It is not yet clear that ‘Indigenous community’ comprises a unitary category in relation to which middle-range generalisations can readily be made.

8. *Defensive agencies.* Evaluations are frequently highly ‘political’ in the sense that there are vested interests in programs being found to be successful (Read & Tilley, 2000; Tilley, 2000). In the UK, police officers sometimes say of their crime prevention initiatives that they are ‘doomed to succeed’, by which they mean that findings of failure are unacceptable. Texts of evaluation reports can be and are massaged to suppress or occlude negative findings and to accentuate positive ones. At worst reports with negative findings are suppressed. The problems are most acute where those involved in the design of a program are also responsible for its evaluation and where further funding turns on positive outcomes. GYFS-NP potentially risks this in that the evaluation is not fully independent of the program.

9. *Implementation failure.* It is one thing to have a plan for an intervention. It is another for the plans to be properly delivered. Apparent negative findings may follow from inadequate implementation of an initiative that could have been effective had it been implemented as intended. In crime prevention implementation failures are common. This is partly because responsibility for community safety is generally assumed to fall to criminal justice agencies, whilst *competency* for much that could control crime lies elsewhere (Laycock, 1996). That is, those generally assumed to be accountable for the prevention of crime, notably the police, are unable directly to influence the conditions giving rise to it. For example, informal control in domestic settings can only be effected by families, the security of car parks by car park owners and managers, and the vulnerability of stores to shop theft by those running the stores. For YSVA in Indigenous communities, it is unlikely that the police have the requisite capability to deliver many of the most promising preventive measures, even if they are able and willing to deliver some. Even where project personnel can deliver the measures

themselves they will need the agreement and collaboration of others. Third parties will, however, need to be mobilised for implementation of most of the measures being designed. This reduces control of implementation and increases the risk of partial or non-implementation of the measures being developed.

10. *Monetisation mysteries.* As already indicated, value for money is a major principle behind rational government decision-making. Yet there are tricky conundrums in the calculation of net effects and accurate monetisation of inputs, outputs and outcomes in community safety initiatives. Net overall effects are important in calculating the value of an initiative. Inadvertent harms may be produced, for example by displacement of offences from one time, place, method, target, offence or offender. Inadvertent benefits may be produced, for example by diffusions of crime prevention benefits by time, place, target, offence or offender. The overall crime effect comprises the direct effect plus the diffusion of benefits effect minus the displacement effects, all calculated and monetised. In practice, tracing and quantifying indirect effect sizes is hugely problematic and so far the only types that have been measured with any sophistication are spatial displacement and diffusion of benefits in relation to programs aiming to reduce domestic burglary. Monetisation of short and long-term non-financial costs and benefits of the effects produced, assuming for a moment that they can be adequately quantified, have tended to use two methods, neither of which is satisfactory. Willingness to pay comprises one: how much would individuals be prepared to pay to reduce or avert a given crime? Willingness to accept comprises the other. How much would individuals accept to suffer a given crime? The first is unsatisfactory in that the ability to pay shapes the potentiality for payment and this will vary by time and community, where acknowledgment of the latter would be objectionable to many on grounds of social justice. The second is unsatisfactory as for some crimes, such as homicide, the sums are likely to be infinite, making calculations of differing returns on spend impossible (Adams, 1995). These difficulties face programs aiming to prevent YSVA in Indigenous communities in a quite acute form. The offences at issue are ones where the primary impacts are non-financial. Indirect diffusion of benefits and displacement effects alongside side-effects that have nothing to do with crime are all possible, albeit that we lack established standard methods for their adequate identification, measurement and monetisation.

11. *Slow burn outcomes.* Many community safety initiatives aiming to reduce risk factors are introduced relatively early in children's lives but cannot be expected to have an impact on the target problem until several years later. As time passes controlling conditions for the purpose of evaluation becomes more and more difficult. Causal attribution is thereby challenging. Yet funding bodies understandably often want robust results much more quickly because of their budgetary cycles. The precise measures to be implemented in GYFS-NP have not yet been fully determined but may relate, for example, to attempting to alter prevailing assumptions about how girls and boys relate to one another, a strategy that will take some years to deliver its benefits.

12. *Initiatives in remote locations.* There are special problems in undertaking fieldwork in remote locations where many Indigenous people live, including one of the sites for GYFS-NP. Some of the problems are logistical; it is expensive and time-consuming to travel to undertake fieldwork. Some of the problems are linguistic; English is the

first language of the research team, whereas a specific Aboriginal language is the first language of the residents. Some of the problems are cultural; non-Indigenous outsiders operate with somewhat different forms, norms and conventions relating to discourse from those prevalent in many Indigenous communities.

13. *Other stuff and secular trends.* A range of well-established sources of error in establishing the responsibility of programs for change face initiatives such as GYFS-NP (Shadish, Cook & Campbell, 2002; Tilley, 2009b). There may be secular trends that explain any observed change. There may be regression towards the mean following a peak that provokes the introduction of the program. There may be selection effects, where program participants comprise an atypical group that will either change anyway or will never change. Communities are not laboratories where conditions can be controlled. They are 'open systems' subject to external events and internal disruption that are liable to interfere with the delivery and relevance of any given program. This is the case for many local programs, where communities are susceptible to disruption from external events and where there are already trends reflecting pre-existing internal dynamics, and where programs may be introduced in response to peaks in rates that fluctuate in a pseudo-random fashion.

Realist evaluation

Realist evaluation takes its inspiration from science. Indeed, it aims to bring scientific method to program evaluation. The Nobel Prize winner for medicine, Peter Medawar, wrote a book on scientific method with the telling title, *The Art of the Soluble* (Medawar, 1967). Being realistic in evaluation means, among other things, focusing on what can be done and setting aside what may seem desirable but is not feasible. The rather obvious precepts that follow are: 'Don't try to measure the immeasurable!' and 'Concentrate your evaluation efforts where robust and useful findings are possible'. In practice this means that no evaluation could, and therefore no evaluation should, aim to measure and monetise all the intended and unintended outcomes that may be produced. It is essential, thus, to be selective and to select from what can be measured and attributed to the program.

Realist evaluation has a particular definition of what comprises science (Pawson & Tilley, 1997, 2009a; Pawson, 2006, 2013; Bhaskar, 1978). Science is concerned with identifying causal potential and causal mechanisms. However science also recognises that few causal mechanisms operate unconditionally. Their activation is contingent on conditions that are conducive to the release of their causal potential. The trick in applied science is to understand where and how to activate the causal potential of interventions to produce positive outcomes that will outweigh any negative ones. Theories are tested and refined over time, using a variety of methods, ultimately producing improved outcomes.

A recent book on the history of cancer treatments neatly illustrates the point being made here. As Mukherjee (2011) shows, cancer treatments have a very long history, rooted in changing theories and producing evolving and contrasting outcome patterns. Progress has consisted in gradually developing better and better understanding of common

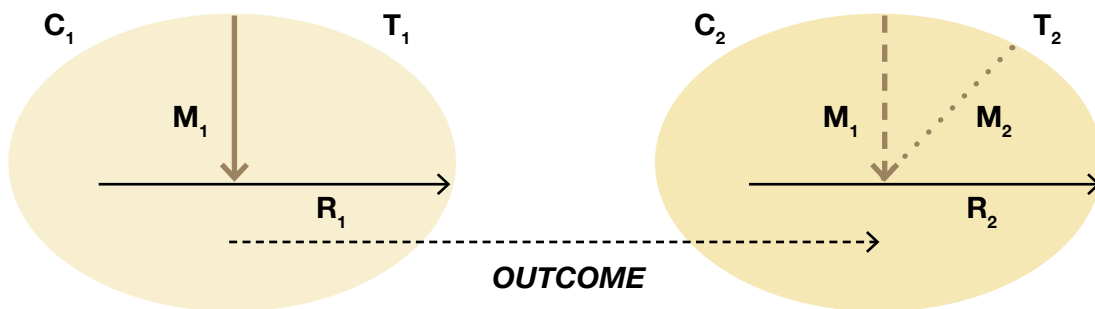
and contrasting features of different cancers together with preventive strategies and treatment regimes that attack different causal pathways and conditions for the cancers to grow and/or change in form and/or spread to further organs. Different theories have been tested using diverse methods, including randomised controlled trials, case comparisons, natural experiments, animal experiments and so on. The methods used have been those that are available and relevant to the theory at hand. Moreover, they have not been conclusive in their findings. Adherents of pet theories have been able to hold on to them! In practice, over time weights of evidence have led to changes in understanding and treatment, leading to lower age-related rates of cancer and better survival times for those who suffer its different forms.

A particular feature of the production of improved cancer prevention and treatment outcomes has been a program of research that has led to a more and more detailed understanding of the causal mechanisms producing cancer and leading to differing patterns of cancer growth amongst those with it. It is this detailed understanding that has paved the way for interventions aiming to pre-empt the production of or inhibit the growth and development of cancerous cells. The story continues, with much of the focus now on understanding the genetics of cancer and the formulation of ways to intervene at that level.

The realist and realistic evaluation planned for GYFS-NP accords with the approach that has been taken in improving the treatment and prevention of cancer. Realistic evaluation emphasises analysis by relevant subgroups furnishing salient variations in context where treatments/interventions activate or deactivate causal mechanisms to generate intended and unintended positive and negative outcome patterns. Figure 1 shows this in a formal way. C describes context. The intervention changes it from Time One (T1) to Time Two (T2). The initial problem pattern (state of affairs such as mortality rate or behaviour pattern, such as YSVA) is referred to as Regularity One (R1). This is generated by the mechanisms (M1) that are activated in the context (C1) furnished at T1. The intervention changes the context and hence the pattern of mechanisms activated. Some will be deactivated (broken M1 in C2) and others will be activated (M2). The new regularities, be they positive or negative or a mix of both, are described as Regularity Two (R2). The Outcome comprises the change in regularities (R2 minus R1) that occur from T1 to T2 as a result of change in context (C2 to C1) from T1 to T2.

There are three important complications not captured in Figure 1. The first is that the processes of change in crime prevention are never fully insulated from unplanned external events that may impinge on the activation or deactivation of relevant causal mechanisms. The second is that there are endogenous sources of instability in community safety; human agents act intentionally and adaptively in the face of change. These mark something of a difference from cancer. Cells do adapt but not intentionally. Moreover, external events do occur but on the whole more slowly. The third is that in the field conditions for community safety multiple mechanisms are activated in complex, overlapping and interacting ways. This is true also for cancer. Thus, Figure 1 is indeed figurative. It represents formal processes but does so by abstracting from much higher levels of complexity on the ground.

Figure 1: The realist evaluation formal framework



Realist evaluation and youth sexual violence and abuse

So what, realistically, will be possible for evaluation in GYFS-NP? The starting point is the theory of initiative and the problem-solving approach, involving some detailed theory-building to inform the data that will be collected and the methods planned for their collection.

The project team has begun by identifying sub-sets of problems, where the participants and generative mechanisms vary quite substantially, although with some overlaps. Sub-problems include, for example:

- a. Peer to peer, casual 'rough sex', which is mostly unreported, to which many of the boys appear to feel entitled, and to which many of the girls appear resigned. Girls submit to the rough sex as a condition for group membership, albeit that taking part in it redefines them as morally reprehensible and thereby unacceptable as long-term partners. The boys take the girls' willingness for granted. The sex generally takes place in secluded open spaces, at night, where boys and girls congregate for recreational purposes. The boys also assume that if the girls socialise with their male peers in those places they know and accept what is expected of them.
- b. Stranger rape offences against adult females, which are rarely committed, normally reported and widely publicised.
- c. Opportunistic youth prostitution for small rewards, mostly provided by adult non-Aboriginal males. The incidents are rarely if ever reported.
- d. Domestic sexual assaults and rapes committed by family members, some of whom are non-resident visitors. This may occur in overcrowded households where boys, girls and adults sleep in the same spaces and where children routinely see underage and adult sexual behaviour, which thereby becomes normalised.
- e. Child sexual assaults and inappropriate proto-sexual behaviour in school, where those involved often cannot be seen because the layout of the school creates

many suitable spaces. The behaviour follows from insufficiently recognised over-sexualised behaviour of children to which no systematic preventive response is currently in place.

There is some evidence in relation to these problems collectively and individually, but it is neither extensive nor direct. Much material comes from community interviews and file reviews, rather than from systematic primary research or the interrogation of detailed administrative records. With regard to the latter, police recorded crime data have been looked at in aggregate, but details of individual incidents have not been available. Rates of sexually transmitted infections have been examined, but again individual case files have not been interrogated. These varying sources complement one another and point in the same direction: that there is a substantial YSVA problem amongst the Indigenous youth in the project's settings and that it takes a variety of forms. The dimensions of the problem and various sub-problems appear to vary considerably between the two communities, though peer to peer sexual abuse seems very common in both. But the direct evidence on the nature and extent of the behaviours is rather limited and the quantitative evidence of their prevalence and incidence is unreliable. Rates of reporting and recording will be low for obvious reasons: some of the actions are at least partly consensual; in some cases the expected personal and familial costs of reporting incidents will likely be unacceptable; and in some cases victims may be too embarrassed to report incidents and may also fear the discomfiture of formal investigation. In addition systematic observations have been made at suspected hot spots, but for obvious reasons these have not specifically concerned sexual behaviour so much as general incivilities and antisocial behaviours. The best evidence suggests that formally available statistical data represent the proverbial tip of a YSVA iceberg, but the actual size or shape of the submerged part of the iceberg cannot be determined with any precision.

In their problem-solving efforts so far the project team has been generating accounts of the nature and causes of the problems, drawing on the available evidence and field visits to sites where the problem behaviours are believed to take place most commonly, including schools, parks and backtracks. Some possible interventions rooted in these accounts have been worked through. The team is also planning additional data gathering to supplement key gaps with what has been available, in order to test emerging hypotheses.

Take peer-to-peer rough sex as an example. Four broad initial prevention activities are proposed, with further analysis needed for three:

1. Enhance formal surveillance and community guardianship at locations found to be at high risk of peer-to-peer sexual assault through a) targeted police patrols, b) reintroduction of community night patrols, and c) installation of CCTV and increased lighting.
2. Design and implement youth based interventions to challenge concerning sexual attitudes and beliefs.
3. Undertake an analysis of youth social behaviour, networks and interactions, and

develop interventions to a) interrupt risky youth social networks and behaviour and b) increase active peer guardianship.

4. Understand and enhance familial supervision and address barriers to guardianship within immediate and extended familial systems.

Here we take just the first of these preventive activities, 'Enhance formal surveillance and community guardianship', focused on injecting police and community patrols. The evidence, theoretical rationale, proposed activity and related provision for evaluation at one of the two project sites are explained.

Evidence

Much of the concerning behaviour involving local youth occurs in public places that are hidden from view, difficult to access for police or other guardians, and provide ready escape routes if anyone tried to intervene. Locations of concern were identified through community interviews, file reviews and direct site observations where the detritus from covert youth behaviour could be observed. Community members reported many specific places to be unsafe, particularly at night, and described community-wide reluctance to intervene in problem incidents at these locations. "Hot spots" were identified on the basis of observed high risk youth activities in combination with low natural guardianship and low formal surveillance.

Rationale

The present design of public spaces presents numerous opportunities for risky and abusive behaviour, as well as significant barriers to formal or informal surveillance and subsequent responses. Given these public spaces have been identified as locations of concern for negative youth activity, general crime, and YSVA specifically, situational strategies to reduce opportunity, increase effort, and increase risk of detection in these locations are warranted.

Increasing formal surveillance and community-level guardianship of these public spaces is aimed generally at altering the existing use of these spaces, and specifically at reducing peer to peer sexual assaults. A range of situational crime prevention activities will be applied to increase effort and risk, in particular (initially) through increasing formal and informal surveillance.

Initial Proposed Activities

- a. Undertakes semi-structured community (adult) interviews focused on understanding barriers to, and opportunities for community guardianship.

- b. Engage with key police staff regarding current patrolling practices, and the purpose and design of targeted police patrols.
- c. Engage with community interest groups, former community patrol volunteers, relevant community service providers, the Regional Council and local church groups regarding the reestablishment of community night patrols.
- d. Engage with Regional Council regarding potential locations for CCTV and increased lighting.

Action Plan

- Consult with local implementation group about proposed activities
- Clarify / review current police practices
- Consult the police regarding targeted patrols
- Engage with relevant stakeholders regarding the establishment of community night patrols
- Consult the cleaning authorities regarding the documentation of rubbish at identified public locations
- Develop and secure formal partnership agreements
- Implement additional research activities
- Analyse results and review intervention plans as necessary
- Implement intervention activities

It is clear that even if the underlying account of the rough sex in public places problem is correct (remembering that much of the evidence for it is anecdotal) and even if the proposed strategy of increasing guardianship is one which could inhibit the unwanted behaviour (by activating mechanisms to do with reducing opportunity by increasing the perceived risk and effort involved in it), enhancing guardianship faces substantial implementation challenges. There is an unstated body of fallible but plausible theory behind the efforts to secure increases in targeted patrols by the police and community members (Jones & Tilley 2004; Ratcliffe, Taniguchi, Groff & Wood, 2011). Moreover, it is possible that any increase in patrols would produce either a) displacement of the youth rough sex to different times and locations or b) diffusions of benefit to other forms of crime and antisocial behaviour taking place in the areas where patrols are planned (for example reductions in robbery, noise, littering, and public drunkenness which are also facilitated by and attracted to secluded places in nearby residential areas known to likely offenders and prospective offenders) or c) diffusions of benefit to times when the patrols were not actually taking place.

In realist terms the theory behind increased patrols is relatively simple:

Context One, with its mechanisms and regularities: The sites comprise easily accessible nearby open spaces for youth, in which covert activities can easily take place, especially at night. These activities include sexual behaviours that boys desire and the girls at least tolerate as the price of group acceptance. The activities involve low effort, low risk and high reward for the boys in particular but also for the girls who are unlikely to make crime reports. There is also no informal social control from significant others in the local community who could show disapprobation. Instead, the group norms of sexual behaviour prevail and it is expected that boys will want sex and girls will be willing to satisfy them. The result is a seldom reported under-age sexual regularity in the particular places. The conditions present also attract other forms of criminal and antisocial behaviour which are also low risk, low effort and high reward. Hence there are non-sexual as well as sexual crime regularities.

Context Two, producing an expected change in mechanisms and hence regularities: Police and community patrols are introduced at those times when illicit sexual behaviour appears to be most common. The change in context means a change in activated causal mechanisms. Real and perceived risks to boys and girls increase, prospective rewards for boys decrease, and effort is increased if the preferred times and places for the sexual behaviour cease to be so readily available. The reduced conduciveness of the context for YSVA also reduces it for other types of crime and antisocial behaviour which also decline as a side-effect, a type of diffusion of benefits. Furthermore because perpetrators are unaware of the patrol times there is a temporal diffusion of benefits beyond the times the areas are patrolled, affecting both sexual and non-sexual crimes, in relation to which prospective offenders overestimate the risks they face. Finally, spatial and temporal displacement may eventually take place as offenders switch the times and places of their (sexual and non-sexual) offences to ones where the risk and effort is lower.

The costs and benefits of the intervention would need to monetise the costs of the patrols and the organisation of the patrols. Against this, monetised benefits from prevented sexual crimes would need to be calculated. Furthermore the monetised costs of any displaced sexual crimes to other times and places would need to be added to the debit side of the equation and the monetised benefits from prevented sexual crimes beyond the time and place of the patrols and of other offences averted as a diffusion of benefits from the targeted patrol would need to be added to the credit side.

The data problems for the evaluation of this small segment of this proposed element of GYFS-NP (one of four types of activity, relating to one of five sub-problem types, in one of two targeted communities) are legion. The patrols are not the only intervention planned by the project team, and there are other initiatives constantly coming and going in the neighbourhood, for example from newly appointed community beat officers and initiatives aimed at other crimes that may impact YSVA. There is little formal reporting and recording of the behaviours, and although the rate of the behaviours is unknown it is likely to be quite small, making tracking real change as against normal

fluctuations difficult to distinguish. It is not possible in principle to know where to look to catch all displacement and diffusion of benefits, and possible candidates will again likely have pseudo-random fluctuating levels. Calculation of the non-financial short and long term benefits and costs involve inexorably contestable assumptions. Producing a precise measure of impact size, costs and benefits is not only likely impossible; it would also be of dubious value if it were too orientated to the specifics of the local setting. A level of abstraction beyond the overly particular is needed if findings are to be more widely usable.

What we have shown here is a tentative theory for one part of GYFS-NP, formulated in realist terms. Its test would comprise a realist evaluation. How, if at all, can it be tested in practice, given all the challenges that face community safety evaluations in general and those relating to YSVA in Indigenous communities in particular? The answer is that those data that can be collected that speak to the theory, imperfect though they are, will be assembled and used. Note that they will not test the whole theory, in particular they will not test those parts of the overall program theory that speak to the implementation of the proposed measures.

Five data sources for this component of the project are being prepared.

- a. Refuse analysis: The suspected high crime areas have regular rubbish collection. The rubbish, of course, comprises litter. It also provides unobtrusive measures of different types of crime and antisocial behaviour: drug paraphernalia indicates drug taking, spent condoms sexual behaviours and empty beer cans alcohol consumption, while the total amount of rubbish comprises an indicator of the total volume of antisocial behaviour. Rubbish collection takes place monthly. In the months leading up to the start of the patrols the collected rubbish will be analysed to obtain a background before measurement. Patrol conditions (police, community, police and community and no patrol) can then be alternated following each clean up and the rubbish next collected then acts as an indicator of changes in illicit usage of the patrolled area. These data will provide an indicator as to whether the types of YSVA and related behaviours are changed. It will also provide some data on possible diffusions of benefit within the patrol target area. Similar observations in nearby areas that are propitious for YSVA will also be made to gauge whether there is displacement or diffusions of benefit to them.
- b. Observations: Controlled observations are planned in the period running up to the interventions. These will involve a checklist indicating who is seen and what they are doing. The observations will be repeated during the different interventions. They will also be undertaken in the hour before and hour after the times patrols are scheduled. Changes and variations in the observed populations and their behaviours comprise a second indicator. This indicator will also provide some evidence of temporal displacement and diffusion of benefits from the patrols.
- c. Interviews: An opportunity sample of young people will be interviewed (using realist methods where the respondent is enjoined to feed back on the theory – see Pawson & Tilley, 1997) as the intervention unfolds to determine whether they have noticed the patrols and if so whether they or their friends have changed their behaviours as a result and if so how. Likewise those involved in the patrols will be interviewed

to elicit their emerging theories about positive and negative consequences of their presence and how these are produced, with a view to testing these practitioner theories if the data that are collected (or can be readily assembled) are suitable.

- d. Police calls for service: These will furnish data on the subset of incidents that come to police attention in the patrol areas and in possible displacement and diffusion of benefits areas. Some may relate to YSVA but only small numbers are expected. The calls for service data will enable comparisons to be made on changes in incident numbers in patrol and non-patrol areas and within patrol areas between patrol and non-patrol times, distinguishing between different patrol modalities.
- e. Patrol records: These will check when patrols are undertaken, what the patrols experienced and what they did.

None of the data sources, a, b, c, d or e, is perfect, but collectively they speak to the theory that has been framed in realist terms. They would not enable a precise effect size to be estimated. They would, however, provide evidence on the direction of the direct and indirect, intended and unintended, outcomes from this segment of the program. Table 2 summarises the broad uses of the data in relation to the realist approach being adopted, and the hypotheses that have been formulated so far.

The detailed analysis, theory-building and ad hoc data collection plans, shown here for targeted patrols, are needed also for all other components of the project if they are to be evaluated realistically. It is a tall order but it marks a way for evidence-based improvement. Where similar types of intervention, rooted in similar theory, are tried in differing sites across the project it may be possible to synthesise findings at a higher level of abstraction to inform policy and practice more widely.

Table 2: **Summary of proposed data uses for realist theory test and elicitation**

	<i>Refuse analysis</i>	<i>Observations</i>	<i>Interviews</i>	<i>Police calls for service</i>	<i>Patrol records</i>
<i>Patrol modalities context outcome variations</i>	*	*	*	*	
<i>Short, medium and long term diffusion of benefits/displacement</i>	*			*	
<i>Implemented measures</i>		*			*
<i>Mechanisms activated</i>		*	*		
<i>Alternative conjectures for possible test with other data collected</i>		*	*		*

Any claim, however, that robust quantified estimated net effects of community safety programs, such those addressing YSVA, including intended and unintended positive and negative effects, monetising both financial and non-financial short and long term costs, would, we think, entail a form of alchemy even though it is clear that they could have potential value to policy-makers and practitioners in allocating their resources. At this stage of development at any rate, as with similar stages in developing treatments for cancer, more modest but nevertheless realistic and important results, emanating from ad hoc opportunist and bespoke data collection and analysis, are needed.

At a much higher level of abstraction some evaluation of the success of the underlying problem-solving methodology is planned. The context is one of 'wicked issues' (Rittel & Webber, 1973): one of unknown extent and severity, that cannot adequately be dealt with using existing standard practices, that has no known simple solution and that falls between or across the responsibilities of several agencies and organisations. YSVA, especially in Indigenous communities, is just such a 'wicked issue'. Wicked issues, the theory goes, can be addressed effectively by a) systematically unpacking the problem attributes in detail and working through possible points of intervention to prevent or ameliorate them, and then b) mobilising relevant agencies and organisations to co-operate and collaborate in delivering changed policies and practices in accordance with that analysis. For YSVA specifically, the public health framework is useful in shaping analysis of the problem and identifying potential points of intervention. Once made aware of the problem and what might be done more effectively to deal with it, well-meaning agencies, organisations and community members can be expected to co-operate in delivering changes in their practices to identify and respond to it. If this overall approach is successful in relation to YSVA, we would expect two measurable outcome patterns. The first would be an initial increase in formally reported YSVA as the problem becomes more widely acknowledged by victims, families and formal agencies and as agencies learn to respond to it more sensitively and effectively, followed by a fall in reports as the overall real rate goes down following effective interventions. The second would be a fall in clinic attendance for sexually transmitted infections amongst young people, given that one source is YSVA, and clinics will, we assume, be attended before, during and after the initiative. These are the broad, expected and measurable outcome patterns.

Conclusions

GYFS-NP provides a unique opportunity to learn about the strengths and challenges associated with the application of Realist Evaluation in the context of an Indigenous community safety project. Realist Evaluation has garnered broad, international respect for its clear focus on building understanding about 'what works for whom in what circumstances'. This combination of questions is appealing. It is essential that administrators of public funds are able to advise government if funded projects achieve intended outcomes. However, for positive results to be maintained or replicated elsewhere, teasing out the contextual elements and the mechanisms that make the outcomes possible is equally essential. Realist Evaluation presents a neat solution to

the seemingly incompatible requirements of contextual specificity and generalisable knowledge. It is hoped that GYFS-NP will both achieve strong safety outcomes for the communities involved, and generate much needed evidence to inform safety initiatives elsewhere. The degree of difficulty of the questions Realist Evaluation seeks to address suggests that successful application will not be easy.

References

- Adams, J. (1995) *Risk*. London: UCL Press.
- Allard, T. (2010). *Understanding and preventing Indigenous offending* (Brief 9, December 2010). Retrieved from Standing Committee of Attorneys-General, Australian Institute of Criminology Indigenous Justice Clearinghouse website <http://www.indigenousjustice.gov.au/briefs/brief009.pdf>
- Allard, T. (2011). Indigenous young people and the justice system: Establishing an evidence base. In A. Stewart, T. Allard & S. Dennison (Eds.) *Evidence Based Policy and Practice in Youth Justice*. Sydney: Federation Press.
- Bhaskar, R. (1978). *A Realist Theory of Science*. London: Verso.
- Byrne, D., & Callaghan, G. (2013). *Complexity Theory and the Social Sciences: The State of the Art*. London: Routledge.
- Cartwright, N., & Hardie, J. (2012). *Evidence-based Policy*. Oxford: Oxford University Press.
- Closing the Gap Clearinghouse (2013). *What works to overcome Indigenous disadvantage: key learnings and gaps in the evidence 2011-12*. Produced for the Closing the Gap Clearinghouse. Canberra: Australian Institute of Health & Welfare. Melbourne: Australian Institute of Family Studies.
- Eck, J., & Spelman, W. (1987). *Problem-Solving: Problem-Oriented Policing in Newport News*. Washington DC: Police Executive Research Forum.
- Eklom, P. (1997). Gearing up against crime: A dynamic framework to help designers keep up with the adaptive criminal in a changing world. *International Journal of Risk, Security and Crime Prevention*, 2, 249-65.
- Farrell, G., & Buckley, A. (1999). Evaluation of a UK police domestic violence unit using repeat victimisation as a performance indicator. *The Howard Journal of Criminal Justice and Crime Prevention*, 38, 42-53.
- Jones, B., & Tilley, N. (2004) *The Impact of High Visibility Patrol on Personal Robbery Hot Spots* (Research Findings 201), London: Home Office.

- Laycock, G. (1996). Rights, roles and responsibilities in the prevention of crime. In T. Bennett (Ed.), *Preventing Crime and Disorder: Targeting Strategies and Responsibilities*. Cambridge, UK: University of Cambridge.
- Laycock, G. (2013). Happy Birthday? *Policing*, 6, 101-107.
- Medawar, P. (1967). *The Art of the Soluble*. London: Methuen.
- Mukherjee, S. (2011). *The Emperor of All Maladies: A Biography of Cancer*. London: Fourth Estate.
- Pawson, R. (2006). *Evidence-Based Policy*. London: Sage.
- Pawson, R. (2013). *The Science of Evaluation*: London: Sage.
- Pawson, R., & Tilley, N. (1997). *Realistic Evaluation*. London: Sage.
- Pawson, R., & Tilley, N. (2009) Realist evaluation. In H. Uwe, A. Polutta & H. Ziegler (Eds) *Evidence-based Practice: Modernising the Knowledge Base of Social Work?* Opladen & Farmington Hills: MI, pp. 151-180.
- Ratcliffe, J., Taniguchi, T., Groff, E. & Wood, J. (2011). The Philadelphia Foot Patrol Experiment: A randomized controlled trial of police patrol effectiveness in violent crime hotspots. *Criminology*, 49, 795–831
- Read, T., & Tilley, N. (2000). *Not Rocket Science?* Crime Reduction Research Series Paper 6. London: Home Office.
- Rittell, H., & Webber, M. (1973). Dilemmas in a general theory of planning. *Policy Sciences*, 4, 155-169.
- Shadish, W., Cook, T., & Campbell, D. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton Mifflin Company.
- Sidebottom, A., & Tilley, N. (2011) Improving problem-oriented policing: The need for a new model? *Crime Prevention and Community Safety*, 13, 79-101.
- Smallbone, S., Marshall, W.L., & Wortley, R. (2008). *Preventing child sexual abuse: Evidence, policy and practice*. Cullompton, Devon: Willan Publishing.
- Smallbone, S., & Rayment-McHugh, S. (2013). Youth sexual violence and abuse: Problems and solutions in the Australian context. *Australian Psychologist*, 48, 3-13.
- Smallbone, S., Rayment-McHugh, S., Crissman, B., & Shumack, D. (2008). Treatment with youth who have committed sexual offences: Extending the reach of systemic interventions through collaborative partnerships. *Clinical Psychologist*, 12, 109-116.

- Smallbone, S., Rayment-McHugh, S., & Smith, D. (2013a). *Preventing youth sexual violence and abuse: Scope, dimensions, and dynamics of the problem at (two sites)*. Unpublished Report.
- Smallbone, S., Rayment-McHugh, S., & Smith, D. (2013b). Youth sexual offending: Context, good-enough lives, and engaging with a wider prevention agenda. *International Journal of Behavioral Consultation and Therapy*, 8, 54-60.
- Tilley, N. (1995). *Thinking about Crime Prevention Performance Indicators*, Crime Prevention and Detection Series Paper 57, London: Home Office.
- Tilley, N. (1996). Demonstration, exemplification, duplication and replication in evaluation research. *Evaluation: The International Journal of Theory, Research and Practice*, 2, 35-50.
- Tilley, N. (2000). The evaluation jungle. In K. Pease & V. McLaren (Eds.), *Crime Prevention: What Works?* London: Institute for Public Policy Research: pp. 115-130.
- Tilley, N. (2004). Applying theory-driven evaluation to the British Crime Reduction Programme: The theories of the programme and of its evaluations, *Criminal Justice*, 4, 255-276.
- Tilley, N. (2009a). What's the 'what' in "what works?" Health, policing and crime prevention.' In J. Knutsson and N. Tilley (Eds.). *Evaluating Crime Reduction Initiatives*. Crime Prevention Studies Vol 24. Monsey, NY: Criminal Justice Press and Cullompton, Devon: Willan, pp. 121-145.
- Tilley, N. (2009b). *Crime Prevention*. London: Routledge.
- Weatherburn, D. (2014). *Arresting Incarceration*. Canberra: Aboriginal Studies Press.