

A theory of change driven approach to evaluating a multi-agency stalking intervention programme

Abstract

Multi agency initiatives as a response to complex crimes and social problems pose several challenges for practitioners and evaluators, in conceptual and operational terms. Challenges for practitioners include working out input requirements to achieve specified interim and final outcomes, and for evaluators it implies measuring not only if interventions work, but how they work, for whom, and under what circumstances. In an attempt to address these challenges, here we present findings from our unique attempt to combine a theory of change driven approach with the realist-inspired EMMIE evaluation framework, of a pilot multi agency stalking intervention partnership in one police force area in England and Wales. The paper makes a significant methodological and empirical contribution to evaluation literature.

Key Words

Multi-agency partnership, evaluation, stalking, theory of change, EMMIE

Introduction

Multi-agency approaches have been the cornerstone of UK policymaking, especially in crime prevention, for at least two decades now, like other European countries, the USA, Canada, and Australia (Balloch and Taylor, 2001; Atkinson, Jones, and Lamont, 2007; Hobson et al., 2021). Over the past decade, multi-agency partnerships have been implemented to respond to child abuse, manage sexual and violent offenders, and tackle cybercrime (Kirkland and Baron, 2015; Smith et al, 2017; Herbert and Bromfield, 2019). Most of these initiatives assume that a multi-agency approach is a more effective, holistic response to crime, as opposed to reliance solely on law enforcement agencies. However, such assumptions are rarely tested. Despite growing public interest in multi-agency partnerships and their potential to address crime-related problems (Vander Laenen and Persac, 2014; Cox, 2019), the evidence base to support the universal belief in the effectiveness of multi-agency approaches is underdeveloped (Chapman, Higgins, and Hales, 2017; Mazerolle et al., 2021).

A theory of change driven approach to evaluating a multi-agency stalking intervention programme

Some studies suggest that there is a little evidential support for multi-agency work being more effective than stand-alone individual agencies (Atkinson et al., 2007; Cleaver et al., 2019; Herbert and Bromfield, 2019). However, they caution that the diversity in definitions, approaches, and contexts must be explored further to identify mechanisms associated with effective implementation. This study fills a gap in the evaluation literature on the effectiveness of multi-agency partnerships by presenting a theory-driven evaluation based on the EMMIE framework (Johnson, Tilley, and Bowers, 2015) of one police force area involved in a pilot project titled 'Multi-agency Stalking Intervention Programme' (MASIP) in England and Wales (E&W) conducted between 2018-2020.

Stalking: a complex social problem

Stalking has been established as a complex and high-harm crime, with heterogenous aetiologies (MacKenzie and James, 2011; Nijdam-Jones et al., 2018). Estimates of prevalence rates of offline and online stalking victimization vary but are universally held to be alarmingly widespread (Björklund et al., 2010; McNamara and Marsil, 2012; Owens, 2016). Whilst no universal definition of stalking yet exists, it is generally conceptualized as 'repeated, unwanted, contact that causes the recipient distress and/or fear' (Korkodeilou, 2016). Thus, it is victim-defined, context specific, and therefore difficult to identify. Some of the behaviours employed by stalking offenders are otherwise perfectly lawful. For example, contacting someone or leaving gifts are welcome in consensual intimate relationships and acquaintanceships. However, when they are unwanted, prolonged, patterned, and accumulate to cause fear, they constitute a criminal offence (Pathé, MacKenzie, and Mullen, 2004).

Stalking manifests in varied and complex ways, making it a difficult crime to respond to (Storey and Hart, 2011). Much of it is hidden from official statistics, as a large proportion of stalking incidents go unreported (Jerath et al, 2022a). Victims who do report often have negative experiences because the police lack a nuanced understanding of stalking behaviour (Holt et al., 2019). The severity of stalking is also generally not appreciated by wider society (McKeon, McEwan, and Luebbers, 2015). It suggests that specialist training around stalking

A theory of change driven approach to evaluating a multi-agency stalking intervention programme

identification and risks is needed for law enforcement to better understand and be able to respond to stalking cases effectively and appropriately (Cho, Hong, and Logan, 2012; Taylor-Dunn, Bowen, and Gilchrist, 2018).

Although legal sanctions might intuitively seem the most effective way to deal with stalking behaviour, clinicians who treat perpetrators claim that these alone do not address their fixation and obsession (Mullen, Pathé, and Purcell, 2000). Other research has indicated that the unwanted contact can escalate, rather than be diminished after legal sanctions (Ostermeyer et al., 2016). Indeed, guidance in E&W¹ no longer advises police cautions be used as a sanction in stalking cases; instead, the Stalking Protection Act (2019) has introduced protection orders as an early intervention mechanism for the police to put protective measures in place for victims who report stalking incidents (Kelly, 2020).

It is well documented that people who stalk have complex needs (HMICFRS & HMCPSI, 2017; McEwan et al., 2017) and clinical interventions form the nascent evidence base into preventing stalking recidivism to date (Wheatley, 2019). However, even after clinical interventions have been delivered, reoffending rates remain stubbornly high at 33-50 per cent (MacKenzie et al., 2009; McEwan et al., 2017). Measurement of stalking reoffending poses greater challenges than other types of offending due to its protracted nature and because subsequent reporting depends on victim satisfaction with the criminal justice response following their initial reporting. Although prosecutions of stalking cases are increasing in the UK, they are doing so from a low baseline (POSTnote, 2018) indicating victim reluctance to engage with the criminal justice system.

¹ It has been advised to exercise care when considering the use of cautions as a police action in practice advice for the police from 2005 (Practice advice on investigating harassment, NCPE 2005). However, it is often the Crown Prosecution Service's decision to advise that a suspect be cautioned, and their advice has, since 2009, increasingly erred towards encouraging cautions to be used sparingly. Cautioning stalkers is inappropriate for several reasons, least not that they deny the victim recourse to a restraining order.

A theory of change driven approach to evaluating a multi-agency stalking intervention programme

Since stalking is a highly nuanced social problem which poses reoffending risks, some advocate multi-agency working to deal with such problems (Rosenbaum, 2002; Pathé et al., 2018). Multi-agency partnerships (hereafter partnerships) appear to be the most viable way of bringing together requisite expertise to effectively manage the risks and harms of stalking.

Conceptual underpinnings of multi-agency working

Unfortunately, the impetus for multi-agency working tends to follow inquiries into tragedies and public crises. For example, the 2003 Laming Report on the Victoria Climbié Inquiry called for improved interagency working in child protection (Sloper, 2004). Similar inquiries into other tragedies have identified one of the main failings to be the inability of agencies to share critical information and to act in a coordinated manner. Presciently, Crawford (1994: 499) noted that multi-agency approaches “are premised on the assumption that failures in criminal justice are ‘system failures’; that along with other local services, criminal justice lacks coordination and systematization.” The criminal justice system can thus be viewed as dysfunctional by the government and the managerial solution to this perceived problem is to create integrated sub-systems that limit duplication of resources and friction between agencies.

The proliferation of crime reduction partnerships in E&W can be broadly categorised into those with a focus on specific crime problems, and those with a focus on high-risk offenders and/or victims. An example of the former is a UK Police Foundation project to reduce recurrent violent crime through multi-agency problem-solving forums (Higgins et al., 2017). Examples of the latter would be Multi-Agency Public Protection Arrangements for sexual and violent adult offenders (Kemshall and Wood, 2013) and Multi-Agency Risk Assessment Conferences to enhance the safety of high-risk domestic violence victims and families (Robinson, 2004).

Several assumptions underlie the accepted wisdom across social and health policy areas that multi-agency working is superior to single agency working. Partnerships are believed to create efficiency gains across the public sector by avoiding duplication of efforts across different

A theory of change driven approach to evaluating a multi-agency stalking intervention programme

agencies (Sloper, 2004). In addition, and *crucially*, they are also believed to work more effectively in enhancing service provision than agencies operating alone (Hester et al., 2017). Reasons given are expedited information sharing, improved referral processes, integration of working practices across organisational boundaries, access to professional expertise, and the ability to present a unified response to service users (Burnett and Appleton, 2004; Robinson, 2004; Sloper, 2004; Berry et al., 2011). Partnerships that focus on managing high risk people can potentially provide a better *quality* of response to service users – due to the synthesis of diverse perspectives on how to respond with the potential to target multiple causal mechanisms driving the problem – and/or a better *quantity* (i.e., dosage) of response to service users (Robinson and Clancy, 2021; Hester et al., 2017; Rosenbaum, 2002). For partnerships responsible for managing high-risk offenders or situations, the gains in effectiveness can be the difference between a calamitous outcome and a less harmful one. Echoing others in the risk management field, Reeves (2013: 41) asserts that there has been a “growing realization that individual risks cannot be managed by focused work undertaken by single agencies working independently to their own aims”.

Role of multi-agency partnerships in managing risk and reducing crime

There are several putative benefits of crime reduction partnerships, and practitioners value them highly (Berry et al., 2011). Furthermore, there are many claims in the literature that multi-agency working delivers greater support and benefits for victims (Lea & Callaghan, 2016), which is a laudable outcome. However, these benefits do not inevitably translate to crime reduction (Chapman et al., 2017). Crime reduction partnerships do not exist in a vacuum. They are embedded within the intricate and interrelated social systems that they are attempting to change. Evaluations of partnerships are consequently, thwarted by the complexity of the social initiatives they seek to investigate.

It is difficult to evaluate crime reduction partnerships because their operations are not static, instead respond to a problem or case as it evolves. Different intervention decisions are taken based on dynamic risk assessments, as risk can fluctuate over time. Partnerships can tailor bespoke *combinations of interventions* to individual cases or problems, depending on the

A theory of change driven approach to evaluating a multi-agency stalking intervention programme

nature of said case. Consequently, there could be multiple treatment effects, which are not easily disentangled in an evaluation. It becomes challenging to isolate individual treatment effects and to divorce them from the multifaceted social contexts in which they exist.

This lack of optimal conditions for traditional experiments is likely the reason for the lack of rigorously designed evaluations of multi-agency partnerships (Rosenbaum, 2002). A notable exception to this is Little et al. (2004) who used a randomised control design to evaluate the Intensive Supervision and Surveillance Programme (ISSP) in Kent, UK. Despite the evaluation team collecting data on the multi-pronged activities undertaken within the ISSP, they did not find any association between the permutation of interventions undertaken on a case and the outcomes. They concluded that the ISSP treatment had a general 'placebo' effect. A major critique of this evaluation would be not acknowledging the role that theory can play in investigating relationships between activities conducted and the outcomes realised. Had a more theory-driven evaluation been completed, we may have a better understanding of what causal processes led to youths on ISSP to reduce their offending frequency when compared to controls.

Theory of change and programme success

A theory of change (ToC) is 'a theory of how and why an initiative works' (Weiss, 1997) which 'can be empirically tested by measuring indicators for every expected step on the hypothesized causal pathway to impact' (De Silva et al, 2014: 4). The term has emerged from evaluation settings characterised by 'uncertainty and emergence' (VanTulder and Keen, 2018). Scholars developed Weiss's (1997) initial definition further to suggest that it needs to provide a framework that explicates the links between activities, outcomes, and the context of the initiative (Connell and Kubisch, 1998), along with the assumptions and risks which may affect the activities (Barnes, Matka, and Sullivan, 2003).

For a ToC to be considered sufficient, it must be plausible, feasible, and testable (Connell and Kubisch, 1998). The plausibility of a ToC is concerned with the logical sense which tie a chain of events leading to an outcome. Feasibility is most important when considering resources,

A theory of change driven approach to evaluating a multi-agency stalking intervention programme

and support, against a political/institutional backdrop which will make an initiative possible to carry out. A ToC driven approach to evaluation provides a promising framework for evaluating complex social interventions, as it helps clarify what should be evaluated, and strengthens the credibility of its conclusions (Connell and Kubisch, 1998).

A ToC is typically developed by evaluators, in collaboration with stakeholders, and adjusted throughout the implementation of an initiative and evaluation via an ongoing awareness of and reflection on exploring change and explaining how it happens (Connell and Kubisch, 1998). ToC models are usually characterised as a logic model which links activities to outcomes and tries to capture causation of change. Without acknowledging the causal mechanisms involved in complex change processes, it is difficult to identify the contributors to a successful outcome (Dyson & Todd, 2010) and therefore different aspects affecting change must be analysed. Blamey & Mackenzie (2007) further note that a ToC should inform an evaluation's purpose, key questions, and consequently, the selection of research methods.

A ToC approach fits neatly within the realist evaluation movement (Pawson & Tilley, 1997). This approach argues that context-mechanism-outcome patterns are observable, and that these generate insight about the conditions under which initiatives might be effective. Whilst realist evaluation has become common in health and social policy evaluation spheres, it has yet to be applied to studying the impacts of multi-agency partnerships. However, it dovetails nicely into what Rosenbaum (2002) suggest for an optimal evaluation of a crime reduction partnership.

The present study attempts to address some of the shortcomings identified in the evaluation literature of previous multi-agency partnerships by adopting a ToC driven approach (Connell and Kubisch, 1998) and using the EMMIE evaluation framework (Johnson et al., 2015), which is inspired by 'realist evaluation' principles. The EMMIE framework is concerned with drawing out not just whether an initiative 'works', but how, for whom, and under what conditions. This framework was designed to capture information (and interactions between) the **E**ffect of an initiative, the **M**echanisms causing the effect, the **M**oderating (i.e., contextual)

A theory of change driven approach to evaluating a multi-agency stalking intervention programme

conditions that activate the mechanisms, crucial information on Implementation and, finally, data on the Economic cost and benefits of delivering the initiative.

We applied this evaluation framework to the Multi-agency Stalking Intervention Programme (MASIP), a pilot project set up in 2018 to reduce the risks and harms associated with stalking across three police areas. In this paper we focus on presenting findings from one pilot site that collected data that was robust enough to enable an attempt at assessing whether the intervention had reduced reoffending². This pilot site operated over a subset of the geographical area covered by Cheshire Police, UK and was called the Integrated Anti-Stalking Unit (IASU). We seek to answer the following research questions:

1. What was the IASU's ToC and through what mechanisms were the interventions intended to work?
2. Was the IASU effective at reducing reoffending and improving victim satisfaction?
3. What are the lessons learned from the implementation of the IASU to guide the setting up of similar partnerships in the future?

Consequently, in this paper we mainly report on three aspects of the EMMIE framework – Effect, Mechanisms, and Implementation³.

Methods

The study setting

The MASIP was a proof-of-concept initiative which aimed to reduce the risk to, and impact of stalking, on victims by developing a partnership intervention model. The initiative sought to

² The other two pilot sites struggled with the data recording and sharing systems across partner agencies in their respective areas and are therefore not being presented in this paper (see Tompson, et al, 2020).

³ There was very little data to test some of the moderators and findings on costs are presented elsewhere (Tompson, et al., 2021)

A theory of change driven approach to evaluating a multi-agency stalking intervention programme

draw expertise and intelligence across the multi-agency spectrum to inform the risk management process associated with managing stalking cases. Key partners common to all sites were police, probation, health, and victim advocates.

Risk management plans were devised for both victims and perpetrators, with victims receiving support from victim advocates throughout the entire case management process. A subset of perpetrators was offered therapeutic interventions to address their obsessive and fixated behaviours when it was deemed to be clinically indicated by health professionals. The stated aims of this project were: to reduce reoffending, increase early intervention to stalking, improve response to victims, improve risk management through enhanced communication between partners, and capture and analyse data to inform national strategy (Belur et al, 2019).

A ToC approach to evaluation using the EMMIE evaluation framework demanded a mixed methods approach as it combines a process evaluation (understanding how the intervention was set up, developed, changed, and what activities were undertaken), with an outcome evaluation (exploring whether the intervention works). Hence, we employed a variety of data collection methods and analytic strategies.

Qualitative analysis

Semi-structured interviews were conducted with six of the core IASU partnership staff and two members of the central project management team at the beginning and end of the initiative. The first wave of interviews explored the aims, barriers, and facilitators for the project, and identified the conceptual elements of the ToC model. The second wave of interviews probed the experiential learning from stakeholders on the benefits and challenges multi-agency working brought to the case management.

Interviews lasted between 10 and 75 minutes and were recorded with the participant's permission. The interviews were anonymized, professionally transcribed, and were coded and thematically analysed using qualitative software NVIVO. Thematic analysis focused on pre-identified themes that answered the research questions and emerging themes were

A theory of change driven approach to evaluating a multi-agency stalking intervention programme

incorporated as the analysis progressed. The data were coded by one member and checked by another member of the evaluation team and analysed jointly. Conclusions were drawn from discussion between the evaluation team members and were refined by feedback from the project stakeholders.

The evaluation team also observed meetings, clinics, and triage processes in the pilot site and held informal discussions with stakeholders to clarify understanding of the processes involved. Documentation of the operating model, provided by the unit, was reviewed. Data collected through various methods were triangulated to deepen reliability - for example, interview data was supported or qualified through our observations and analysis of documents.

The combination of qualitative methods used over the fieldwork informed the generation of *a priori* hypotheses (see Jamal et al., 2015). Hypotheses related to the *mechanisms* (the way in which IASU exerted its effects), the *moderators* (the pre-existing contextual conditions) and the *implementation* (conditions introduced by IASU and the evaluation). These were generated before quantitative data were collected, to protect against subjectivity.

Quantitative analysis

Comprehensive quantitative partnership data was provided by IASU. This contained information on the perpetrator and victim (including whether they had been known to any public services previously, such as probation or mental health), details of the case, details of what interventions were used during the case management, and a field that stated whether further stalking behaviour had been reported at three-monthly intervals. The stalking behaviour could be reported by victims to the police or to their victim advocate, by probation, or could come to the unit's attention via other means. Consequently, it is the most comprehensive measure of stalking reoffending, of which we are aware, in the literature, made possible by the cohesive multi-agency partnership.

We defined reoffending as additional stalking behaviour within six months of the original offence coming to the attention of the IASU. This six-month period accords with research that

A theory of change driven approach to evaluating a multi-agency stalking intervention programme

defines continuation of stalking behaviour within a six-month period as constituting the same 'episode' or offence (McEwan, Mullen, and MacKenzie, 2009). However, due to the severely restricted evaluation timescales (17-months with no follow-up period funded) we deviate from research that uses the conclusion of a health intervention as the starting point for measurement. This means that we are liable to overestimate reoffending since we start measuring at a much earlier point (the antecedent offence), and before any kind of intervention is done.

Secondary criminal justice outcome data on stalking offences was secured from a Police force judged to be most like Cheshire⁴. This was for the period of August 2016-December 2019, inclusive, and hence does not include the full study period.

Descriptive analysis was undertaken of partnership data and the secondary criminal justice outcome data, including chi-square and Fisher's exact tests. Logistic regression was run with the reoffending outcome (captured in the partnership data) as the dependent variable and intervention types by the IASU as independent variables. Due to the sample size being small (n=71), we ran two versions of the following logistic regression model; one standard and one with Firth's (1993) method, that acts in a similar way to penalized likelihood to correct for biases associated with small samples. Prior to running the models, we first dichotomized the independent variables so that they were mutually exclusive.

Findings

Cases were referred into the IASU by a range of agencies and were initially screened and triaged. Regular internal meetings to discuss cases were facilitated by the co-location of core

⁴ We are required to keep the force selected anonymous for data protection reasons. Forces were selected as comparators based on their similarity to the three MASIP pilot sites on the following criteria: 1) Most similar forces (produced by the Home Office); 2) Reported stalking offences trend trajectory, and 3) Police and Crime Commissioner (PCC) funded services (i.e., where stalking was mentioned as a priority in the PCC's strategy, but usually subsumed within domestic abuse).

A theory of change driven approach to evaluating a multi-agency stalking intervention programme

staff, and these supported the functioning of the fortnightly partnership meeting. The partnership meeting was where wider partners (e.g., Probation, Health, Social Services) were invited to risk assess specific cases, and bespoke advice was dispensed to referring agencies (e.g., investigative advice to the police). These partnership meetings were described by one interviewee as,

“In the multiagency meeting to assess the risk, decisions are made. We share information first and then we go through our clinic assessment form methodically, so we look at the risk and then we assess that together, so the partners all put their opinion in of whether they think first of all is it a consensus, is it stalking, what the risks are, what the typology is. We all input. The partners don’t just share information, they also input on our decision-making regarding risk and that we share that risk.” (CSs03, Social Services)

Complementary tasks undertaken at these meetings included signposting to appropriate agencies and the creation of safeguarding plans for victims and perpetrators. The IASU sought to identify the most appropriate interventions to manage the risk in a case, which often integrated legal sanctions with inputs from mental health services. Victims and perpetrators were hence supported directly and indirectly within this partnership model.

We present our findings in response to the three questions guiding this research.

What was IASU’s ToC and mechanisms for change?

Unusually for a multi-agency initiative (Rosenbaum, 2002), stakeholders were encouraged during the project set-up phase to develop their individual ToC by the project managers. IASU stakeholders were encouraged to present their initial theory of change and we used this to initially populate our input-output-outcome-impact model for IASU. However, based on our observations and interviews with stakeholders, we refined their ToC to depict how the MASIP model was intended to work to achieve articulated outcomes (see Figure 1). Our experience indicated that the partnership found it difficult to succinctly articulate the input-output-outcome and impact model we supplied. This is because practitioners tend to think only in terms of what actions they undertake and what results they hope to achieve. Thus, boundaries between what Rogers (2014) would traditionally call inputs, outputs, outcomes,

A theory of change driven approach to evaluating a multi-agency stalking intervention programme

and impact dimensions are often ambiguous. To elaborate, when asked what would be considered successful outcomes, interviewees made a clear distinction between what they considered to be successful outcomes for the unit, as well as what they would consider successes for their own agency. For instance, perceptions of success ranged from “generating masses of data for the public good” (CPo2, Police), to leading the perpetrator towards self-regulating their behaviour, as one interviewee said,

“For them to have worked, with our guidance to a place where they're in employment and they haven't been in two years, they're not drinking alcohol anymore, using more appropriate coping strategies to manage emotions, being aware of their emotions and recognising when they're experiencing different things and being aware of their own triggers or what might cause them to reoffend, I think that might be considered a success”. (CH14, Health/Outreach)

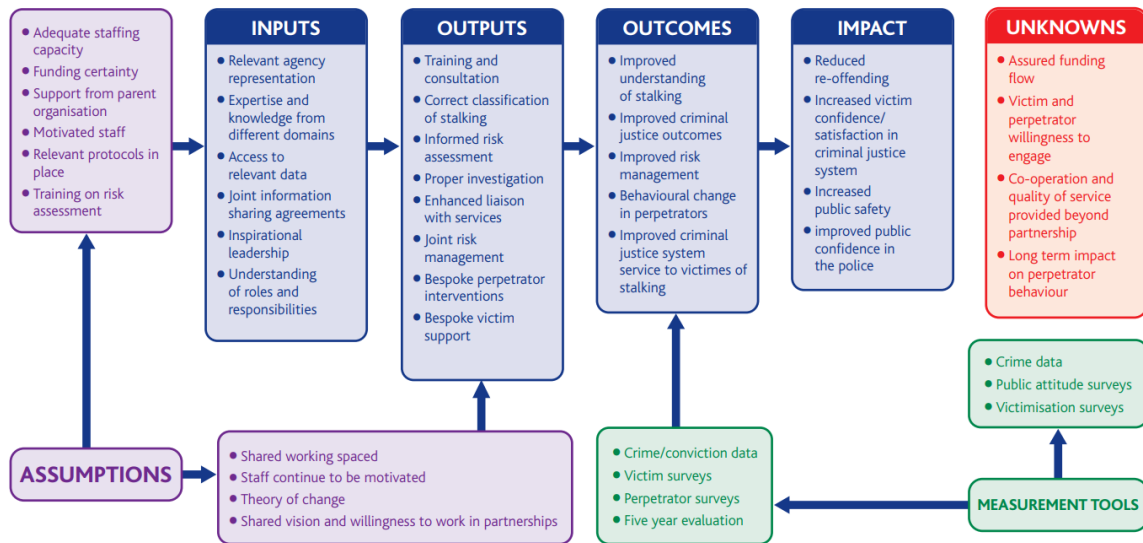
The notion of successful outcomes included correct classification of stalking by the police instead of just harassment, as well as supporting and empowering victims (CVa16, Victim advocate); upskilling and raising awareness in other relevant agencies (CPo2, Police); and taking a holistic approach to risk management,

“We had a blank page to start with. Then that expertise that those individuals can bring in that multidisciplinary setting. We have got very senior practitioners across all of those disciplines who are able to bring a really expert perspective to address the problem” (SLT003, Central project management)

Based on our understanding of the IASU's workflow processes from the stakeholder interviews (see Figure 2) and the literature on multi-agency partnerships dealing with high-risk cases (Robinson, 2004), we hypothesised that three possible causal mechanisms could bring about the first aim of MASIP, i.e. reduction in reoffending. The first of these was: improved classification of stalking cases and better investigation, aided by informed case review, would lead to higher convictions, thus ensuring the offender is restricted in reoffending for the time they are in custody or under court mandated restrictions. Steps on the path to increased convictions were identified by our interviewees as desirable interim outcomes, with knowledge and information sharing being one of the most important of them. As this interviewee said, “a lot of what we wanted to do as a unit was around spreading knowledge, sharing expertise, education”. (CPo1, Police).

A theory of change driven approach to evaluating a multi-agency stalking intervention programme

Figure 1: Programme Theory of Change for IASU/MASIP



Stakeholders indicated their awareness of the fact that conviction is unlikely to address the underlying obsessive fixation underlying stalking behaviour, however, stakeholders representing different agencies highlighted the importance of improved conviction for different reasons: police - because increased convictions would improve public confidence in the criminal justice system and also because it restricts opportunity to commit offences while incarcerated; probation - because it provides them the opportunity to supervise and work closely with perpetrators; health - because it allows for the opportunity for health professionals to assess whether the perpetrator would benefit from health interventions and/or for the court to mandate conditions or treatment that might address the fixated behaviour, provided the perpetrator is willing to engage with it. Finally, victim advocates – because it provides victims with reassurance that their complaint is being taken seriously and because it possibly restricts the offender from offending whilst incarcerated.

The second mechanism proposed was that bespoke health and other interventions designed to address the perpetrator’s needs would lead to change in offender behaviour, ultimately leading to a reduction in reoffending. One of the biggest ‘successes’ of the project was the

A theory of change driven approach to evaluating a multi-agency stalking intervention programme

delivery of bespoke treatments designed for individual stalkers. As one stakeholder concluded,

“But even within types of stalkers, there is variance and differentiation and diversity... you need to understand what’s driving the behaviour for that individual at that time, rather than thinking, ‘Oh well, they’re this type of stalker, so therefore this is the motivation.’ That’s not the case. And also, therapeutically, we’ve had to adapt and modify how we’ve worked with individuals, because of their competencies and lack of competencies, and skills and lack of skills, and underlying intelligence levels; we’ve had to adapt and accommodate and modify because of their learning styles, which is right.” (CH3, Health)

Finally, the third mechanism suggested was that multi-agency working, by its very nature of information sharing and working together will lead to a more effective and informed risk management plan and its execution, thus leading to reduction in opportunities to reoffend by being able to see both sides of the interaction. As one interviewee said,

“So, working in the unit where we have had information being shared from both the perpetrator and victim allows professionals to understand the whole story, or a lot of the story compared to what both individuals are saying, that the story may be in the middle somewhere. It just keeps adding pieces to that jigsaw, and you're starting to be able to see what's happening and being curious to find out what those missing pieces are” (CH15, Health/Outreach).

Another interviewee took this concept further and said that by the very nature of the multi-agency partnership, the stalking problem profile was raised from being a criminal issue to one of wider social significance,

“It’s beyond that task, that incident, that crime that you’ve got. The wider social consequences, economic consequences, for the victim, the perpetrator, and all those connected to them couldn’t be further from most responding officers’ minds, when in actual fact, what we start to do through our intervention, through our practice, through our advice, through agencies working like this, it brings it into the narrative, doesn’t it, this idea that this is a problem beyond the criminal justice system.” (CPo1, Police).

Knowledge sharing and multi-agency working also helped individual stakeholders change their perspective from being focused on their individual agency outcomes to a wider understanding, as this interviewee said, “I’ve gone from being very victim focused to being

A theory of change driven approach to evaluating a multi-agency stalking intervention programme

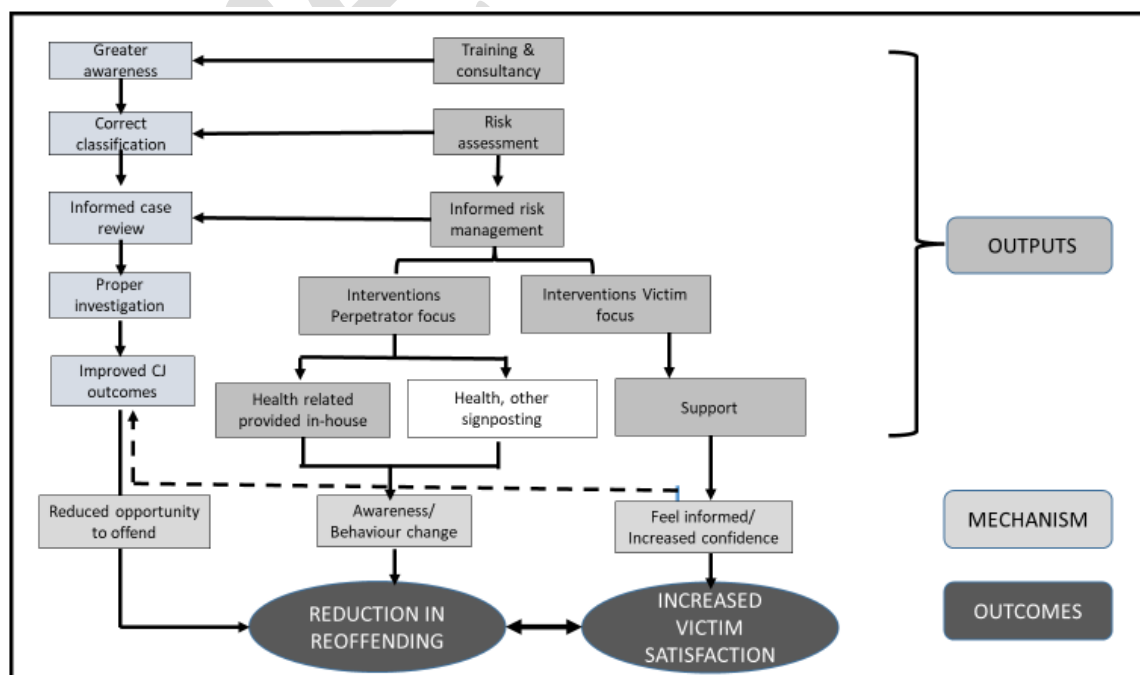
holistically focused". (CVA16, Victim Advocate). Furthermore, as one interviewee from the central project management team explained the advantages of multi-agency working was,

"I think there is also something about having an awareness of our differences and letting go of that preciousness about it, really. Being able to recognise that there is more than one way of resolving all of these issues and that nobody is necessarily right and nobody is necessarily wrong." (SLT001, Project Team)

It was thus clear that some of the articulated aims of the project were but essential steps in achieving the two ultimate goals of reducing reoffending and improving victim satisfaction.

We also identified two mechanisms that were activated to achieve the second intended objective, i.e., improved victim satisfaction. The first one was that bespoke support provided to victims in the form of emotional support, advice on how to collect evidence, legal processes, and risk management techniques would lead to an increase in satisfaction. Secondly, a reduction in reoffending through the three causal mechanisms identified above would ultimately lead to a reduction in experience of re-victimisation and consequently, increase in victim satisfaction and a reduction in fear of revictimization. This is a more of a circular mechanism whose impact is likely to be more evident in the long term.

Figure 2: Process Map for IASU/MASIP



Was the IASU effective at reducing reoffending and improving victim satisfaction?

Over the data period September 2018 – December 2019 the partnership classified 103 stalking cases as those that they believed that they could add value to. This referred to whether the unit would provide some actual support to either the investigation or the risk management process as well as cases where the team had the capacity and capability to provide support, therapy, or services inhouse.

The first mechanism proposed to reduce reoffending was higher convictions of perpetrators, which was tested by comparing the charge and caution rates for Cheshire with a comparator police force in the period August 2016 - December 2019. Detailed results are published elsewhere (Tompson, et al, 2020), but in brief, this revealed that Cheshire (IASU) brought charges or summons in 17.7% of stalking cases, compared to 14.6% in the comparator force. Since charges take a considerable amount of time to eventuate and the data period only captured the beginning phase of MASIP, the true effects of IASU to increase convictions of perpetrators is likely underestimated. However, these nascent data suggest that higher convictions were indeed being achieved by IASU when compared to a similar force area.

To test the second mechanism, whereby reoffending is hypothesised to be reduced by the bespoke health and other interventions delivered by IASU we turned to the partnership data⁵. Of the 71 cases where six months or more had elapsed since the unit began monitoring the case, there was evidence that 30 stalkers, not all of whom received a health intervention, continued their stalking behaviour. This equated to 42.3 per cent of the sample reoffending. As mentioned previously, this is likely an inflation of the true reoffending rate because this is optimally calculated from when the health intervention concludes, and this date was not

⁵ As police agencies outside of the pilot sites were not devoting additional resources to the accurate recording of stalking, we could not use these data to test research questions associated with reoffending.

A theory of change driven approach to evaluating a multi-agency stalking intervention programme

captured by IASU in the tight project timescales. This reoffending rate is within the range reported in similar studies (10-60 per cent; see Appendix).

The full breakdown of reoffending can be seen in Table 1⁶. Visual inspection of this shows that 20.4 per cent of stalkers had reoffended by three months; 42.3 per cent by six months; 47.7 per cent by nine months and 58.3 per cent by twelve months.

Table 1 – Quarterly breakdown of reoffending

Reoffended	< 3 months		< 6 months		< 9 months		< 12 months	
	n	per cent	n	per cent	n	per cent	n	per cent
Not reached time threshold	10		32		59		79	
No	74	79.6	41	57.7	23	52.3	10	41.7
Yes	19	20.4	30	42.3	21	47.7	14	58.3
Total	103		103		103		103	

To consider the impact of the third mechanism – that multi-agency working would lead to more effective risk management – we conducted separate analysis on the reoffending outcome data. Of the 44 cases where more than six and less than nine months had elapsed since the IASU had been monitoring the case, there was evidence that stalking behaviour had recurred (i.e., after a break of 6 months) to the same victim in 4.5 per cent of cases (n=2). This is lower than the range that is reported in research that has studied recurrence following a psychological intervention (9.5 per cent - 40 per cent, see Appendix), and in these two cases perpetrators had not received a health (psychological) intervention from the IASU. This suggests that risk management interventions directed towards victims – for example the safety planning they received from the Victim Advocates that was informed by the multi-

⁶ Re-offending was counted once only in Table 1. Put otherwise, if reoffending had (say) been recorded within the first and third quarters it would only count as one instance in Table 1 at the earliest occurrence.

A theory of change driven approach to evaluating a multi-agency stalking intervention programme

agency risk assessment - may have indirectly decreased their chances of re-victimization. Although with a small sample such as the one here it is naturally prudent to advise caution on the reliability of these results.

Table 2 provides a breakdown of the 71 cases where enough time had elapsed to judge recidivism at six months post IASU involvement, by the intervention type that was used in the case. These are not mutually exclusive categories. All victims received safeguarding, so this is not considered an intervention per se. Of those perpetrators who received a direct health intervention, 17.6 per cent reoffended. This is at the low end of the range when compared to reoffending rates in studies where the perpetrators have received a health intervention with the measurement caveat expressed previously (0 - 52.9 per cent, see Appendix), and suggests that the health interventions delivered in the pilot site were modestly successful⁷.

Table 2 - Intervention type by whether there was any evidence of reoffending

Intervention type	Yes	No	Total
Comprehensive risk assessment	4	6	10
Perpetrator direct intervention	3	14	17
Perpetrator health signposting	4	1	5
Perpetrator 3 rd party consultation	21	26	47
Victim direct intervention	10	11	21
Victim target hardening	2	5	7
Victim 3 rd party consultation	24	29	53

We present the results of the logistic regression in Table 3, with the coefficients converted into odds ratios to assist interpretation along with 95 per cent confidence intervals. Visual inspection of Table 3 reveals that none of the independent variables, which represent various

⁷ It is important to note that there were myriad reasons why direct health interventions were not always delivered. A perpetrator could be unreceptive to treatment, could be assessed and found not to be suitable for treatment, or could move out of the pilot area or become deceased.

A theory of change driven approach to evaluating a multi-agency stalking intervention programme

intensities of multi-agency working with perpetrators and victims, are statistically significant. Applying the Firth method confirmed these results⁸. Whilst disappointing, these results are unsurprising in the context of the sample being severely underpowered.

Table 3 - Logistic regression; reoffending at 6 months as the dependent variable (n=71), AIC: 102.53.

Intercept	0.86	0.77	0.18	3.97	0.85
Perpetrator direct intervention	0.74	0.85	0.12	4.51	0.74
Perpetrator indirect intervention	0.96	0.81	0.20	5.01	0.96
Perpetrator combination intervention	1.94	0.91	0.37	10.98	0.44
Victim direct intervention	0.48	0.75	0.09	2.22	0.36
Victim indirect intervention	0.41	0.80	0.11	1.38	0.16
Victim combination intervention	2.20	0.63	0.52	10.46	0.30

With regards to measuring the impact of IASU on victim satisfaction, we intended to capture victim satisfaction - via a bespoke exit survey that we co-designed in consultation with victim advocates involved in the programme. However, this was not used by the IASU for no specific reason, except that perhaps they did not realise the importance of collecting this data. Thus, we had to rely on qualitative data in the form of interviews with victims and stakeholders to assess the impact of the partnership on victim satisfaction (reported in Jerath, et al., 2022a).

In addition, stakeholders reported that they had received plentiful positive feedback from victims, and shared letters and cards thanking the unit and the advocate specifically with the evaluation team. The value of having a multi-agency set-up seemed to benefit victims with regards to their safeguarding and risk management.

“I think it’s almost as if it’s become crucial for us to have it because it’s definitely filled a gap within services for victims, and obviously in my close relationship with the

⁸ Results available on request from the corresponding author.

A theory of change driven approach to evaluating a multi-agency stalking intervention programme

IDVAs⁹ and other partners we've had, I'm working with them when they come across victims and having that wraparound support service for them has really made so many differences in people's lives." (CVa16, Victim Advocate)

Overall, interviews with stakeholders and victims indicated that whilst they were satisfied with the support they received from the partnership, they were less satisfied with the criminal justice system.

What implementation lessons were learned to guide future partnerships?

From our qualitative fieldwork we identified factors necessary for the effective and efficient working of multi-agency partnerships and the successful achievement of intended outcomes. These are intended to guide models of partnership working similar to the one reported herein. Some of these, such as the importance of data sharing agreements, are specific to the E&W context and may not be applicable to partnerships elsewhere.

Facilitating factors are grouped into three sets: essential preconditions; essential for continued performance; and essential for ensuring sustainability. Most factors are within the control of partnership agencies, but some are dependent on external factors, outside the remit of partnerships.

Essential preconditions

The first essential precondition is full organisation and individual stakeholder commitment to the partnership work. Stakeholders interviewed insisted that all key players involved in the partnership ought to be motivated and willing to work as a team to be effective. As this interviewee said,

"It's the right people being involved. Not somebody being sent to a meeting because that's your job to go to meetings. We want people who are interested, passionate,

⁹ Independent Domestic Violence Advocates

A theory of change driven approach to evaluating a multi-agency stalking intervention programme

willing to learn and engage and actually they see the benefit in what this does for them...They come along, galvanise the catalyst for change. These are the people that drive the cultural attitudes, the peers..." (CPo1, Police)

The willingness to work together is also predicated on support from senior leaders in the partner organizations. The setting up of the MASIP project augmented funding and resources to the already existing unit in Cheshire, thus gaining the backing of senior leaders in the respective organisations. However, stakeholders made it clear that substantial impetus and the leadership required for joint working came initially from individuals within relevant agencies. Furthermore, there was acceptance that without the support of the senior leadership team, partnership working would be very difficult.

Secondly, our observations indicated the importance of having information sharing agreements in place prior to or at the start of the partnership. These often take longer than expected and can act as an impediment to integrated working. Fortunately, IASU had these in place due to the pre-existing unit. This helped pave the way for smoother and more efficient partnership meetings and setting up databases to monitor the progress and outcome of cases over the course of the pilot.

Thirdly, the evidence indicated that joint training on stalking and how to work in partnerships would be useful. Partnership members require a shared language and shared understanding of the problem to conduct joint risk assessments. Stakeholders rated the two-day Stalking Risk Profile (SRP) training they received early in the project as crucial. As this interviewee said, "I think MASIP has also given us the opportunity to go on training we would never have gone on before." (CSs3, Social Service). When asked, stakeholders said they had never received training on partnership working, despite some of them working in partnerships for over 20 years. They reflected that they would have valued such training, including understanding each other's databases, key performance indicators, and chain of command.

Essential for operational continuity

At its core, adequate resourcing and time are essential for the project set up and operational phases. Stakeholders mentioned that recruiting and training suitable personnel took a

A theory of change driven approach to evaluating a multi-agency stalking intervention programme

considerable amount of time. Furthermore, individual stakeholders were often overworked, or needed more support in terms of resources. For example, petty cash needed to be made available so that the IASU victim advocate did not use personal funds for refreshments when meeting with victims in safe public spaces to gain their confidence. Sustained drain on specific agency or individual resources, financial or otherwise, can prove detrimental to the success of the partnership work.

Secondly, co-location and a shared working space were highly valued by the IASU stakeholders so that partners could easily access relevant expertise from partner agencies and take joint decisions.

“There are definitely processes that developed along the way that we've all had to learn and actually being located in the same office space I think has been invaluable in making sure that everyone's roles are contributing to the same overall vision.” (CH14, Health/Outreach).

Furthermore, the importance of building relations of trust between partners cannot be overstated. This invariably takes time. IASU had mature working processes whereby close working interpersonal relationships had been nurtured over time and appropriate governance structures and mechanisms were in place to resolve disagreements. As one stakeholder, external to the core IASU, but part of the wider multi-agency consultation said about the IASU,

“I think the partnership works quite well because you have no hierarchy in the meeting which I think is really good, there's no fear to speak or contradict, which I think is a really positive... I think there is the freedom to disagree, there's the freedom to talk something through and I think by doing that you can come to a much better conclusion of what the risks are, how we can try and manage risk, how we can try and protect people or advise other agencies we are not happy with this, you need to look again.” (CSs3, Social Service)

Another interviewee commented on the result of trusting relations built over time on the working of the IASU,

“We are integrated, we sit side by side, we share competency. When we sit around the table, we don't talk as psychologist, outreach practitioner, support office, and ISAC, we speak as a collective.” (CPo4, Police)

A theory of change driven approach to evaluating a multi-agency stalking intervention programme

And finally, although it sounds common-sensical, the willingness of perpetrators and victims to work with the partnership is vital. In the absence of engagement from perpetrators and victims the partnership has little chance of effecting behavioural change. Trust and confidence from victims, especially, is not a given (Robinson, 2004). Similarly, stakeholders acknowledged that reducing the obsessive compulsive thought patterns that impelled stalking behaviour through therapeutic interventions was dependent on perpetrators wanting to engage with the process, as the team did not have any authority to compel them to engage with therapy.

Essential for ensuring sustainability

At its core, adequate and assured funding is the lifeblood for the success of a partnership. Funding uncertainty can undermine treatment continuity in therapeutic interventions. The IASU stakeholders reported experiencing funding uncertainty which led to a loss of some personnel, and hence expertise.

Relatedly, personnel stability in the partnership is vital. Acquiring knowledge and expertise in dealing with a complex crime such as stalking requires considerable investment by individuals in terms of time and resources. Furthermore, evidence indicated that it is also essential that there is some constancy or proper handing over of senior management roles within various partner agencies, as this interviewee expressed their frustration,

“All the changes in their management and leadership haven't helped either; there's been no real ownership, a lot of empty promises from them about support and stuff like that. They have done some stuff, but it's been, I think, more about us satisfying their needs rather than them helping this project succeed.” (CPo2, Police)

Discussion

In this study we report a theory-driven evaluation of an English multi-agency partnership to prevent the harm associated with stalking. A ToC approach has several purposes from contributing to resourcing and planning, describing the process, aiding with monitoring and evaluation, and identifying feed-back loops (Stein & Valters, 2012). In our evaluation we found the ToC approach useful for two reasons: first, it helped guide practitioners to establish

A theory of change driven approach to evaluating a multi-agency stalking intervention programme

how programme activities, supported by appropriate resources, could be aligned with intended outcomes via clearly articulated mechanisms. Second, it supported a richer contribution to the evidence base as it enabled us to generate and test *a priori* hypotheses to better understand how or why interventions may work (or not).

The ToC elicited from stakeholders in the first phase of the process evaluation enabled the generation of research questions and hypotheses that were subsequently tested with qualitative and quantitative data collected. The evaluation's mixed methods approach adopted the EMMIE framework (Johnson et al., 2015) which prompts that we not only ask 'what works?', but also 'how it works, for whom, and in what circumstances?'. This requires attention to the mechanisms and the circumstances in which those mechanisms are activated, neutralised, or deactivated.

We set out to answer three research questions: how was the IASU envisaged to work? Did it work (in reducing reoffending and improving victim satisfaction)? And what implementation issues were identified? Our initial fieldwork indicated that since the IASU had been in operation for some time in another form before the start of MASIP, they had an established working ToC, which just needed refining. Clear visual depiction of a ToC can be a powerful "tool for describing the key components of an intervention and how it impacts on outcomes" (De Silva et al., 2014: 10), and it can help other researchers and policy makers understand how a complex intervention works. However, visually depicting complex multi-agency initiatives where several parts occur simultaneously or sequentially can be 'a mess of criss-crossing boxes and arrows' (Funnell and Rogers, 2011). IASU stakeholders clearly had an intuitive version of their ToC which provided them with a kind of operational clarity, but they struggled to try and depict it visually. The evaluation team worked with the stakeholders in presenting a simplified and linear process diagramme (Wilkinson et al, 2021) which linked different activities to outcomes (Figure 2).

The ToC enabled the identification of several mechanisms to achieve the intended outcomes of reducing reoffending and improving victim satisfaction. We subsequently used the insight provided by the ToC to set out *a priori* hypotheses to be tested with the quantitative data.

A theory of change driven approach to evaluating a multi-agency stalking intervention programme

This enabled us to design an evaluation to understand both outcomes and causal mechanisms (Breuer, et al, 2015). We specified the data trends we expected to see if the mechanisms were working as intended in advance of the quantitative data analysis. (Belur, et al., 2019). This guarded against loss of independence (i.e., working closely with the stakeholders) and subjectively selecting results that biased the findings of the evaluation.

Overall, the hypotheses regarding whether the IASU partnership reduced reoffending are not strongly supported. We found only partial support for the hypothesis that improved investigation would lead to higher convictions in stalking cases. However, given that demonstrating evidence of effect would have been extremely challenging in the short evaluation period, the reoffending rates emanating from the IASU indicate that there is cause for optimism. For example, 17.6 per cent of all perpetrators receiving a psychological intervention reoffended by six months, which compares favourably to commensurate rates reported in studies of other psychological interventions (10-52.9%, see appendix). The broader reoffending rate (for all perpetrators) was 42.3% at six months, which again falls within the range reported in similar studies (10-60%, see appendix). Given that measurement of reoffending began at the point the case came to the attention of the IASU – rather than at the completion of treatment as used in other studies - it is likely that the IASU reoffending rates are inflated in relation to comparators. This is further compounded by the measure of stalking reoffending being more comprehensive than other studies, due to the excellent communication between the victim advocate and the victims. Hence, the effects of IASU are likely underestimated here.

Further, the uncertain funding situation to support continued operation beyond the initial 18 months prompted ethical concerns vis-à-vis treatment completion, which stakeholders believed undermined the optimal delivery of the health interventions. Additionally, it is worth stressing that not every offender is willing to engage in treatment. Although, qualitative analysis of interviews with victims and perpetrators reported elsewhere (see Jerath, et al, 2022a Jerath, et al, 2022b) revealed that there is reason to believe that the interventions delivered by the partnership certainly benefitted some participants in the programme.

A theory of change driven approach to evaluating a multi-agency stalking intervention programme

Crime reduction interventions often fail to have the intended effect, but it is not always clear whether this is attributable to theory or implementation failure. Implementation failure is particularly a threat in social programmes that depend on a protracted and complicated chain of events to occur before an intervention can be considered completed (Bowers et al., 2018). Our evaluation helped identify essential preconditions to facilitate partnership work which included serious commitment and willingness from senior leadership in organisations and stakeholders to work together; information sharing agreements in place; and possibly some joint training to facilitate partnership work. Conditions to support smooth partnership working included adequate resourcing, co-location where possible, adequate time to build trust between partners and, finally, willingness of perpetrators and victims to engage with the partnership. Finally, for partnerships to be sustainable, assured funding over a sustained period is key, as is stability of personnel to ensure co-produced knowledge remains protected. Support for complex social interventions and partnership-based initiatives like IASU to succeed require continued patronage and support from government, policymakers, and criminal justice agencies as well as the health sector.

The pros and cons of a ToC driven evaluation

The literature reports several key challenges with taking a ToC evaluation approach. Here, we discuss some of these challenges and how they were addressed in our evaluation. The foremost challenge arises from the fact that complex initiatives often have vague designs. Consensus of long-term outcomes can be straightforward, however stakeholders can struggle to specify interim outcomes and activities and how they fit in a causal process (Connell and Kubisch, 1998). Despite the IASU having a good operational grasp of how their partnership intended to achieve its aims, we experienced this challenge in articulating causal processes.

Compounding this, it is not unusual for partner stakeholders to have conflicting ideas of what outputs might best lead to desirable outcomes. For example, if the overall goal of an initiative is to reduce reoffending, the police may approach this through more detection and incarceration, while probation and health may think that receiving appropriate treatment would lead to desistance. Agreeing on a ToC, with shared outcomes and impact can be

A theory of change driven approach to evaluating a multi-agency stalking intervention programme

difficult in partnerships with multiple stakeholders, However, given the fact that the IASU already had built trust between partners, partnership working was made easier because of shared goals within the team.

The process of designing the ToC to the satisfaction of all partners was iterative and intensive, and involved a workshop with stakeholders and myriad other fieldwork activities (e.g., observations, interviews). Hence, it was not merely a 'box ticking exercise' as De Silva et al (2014) caution against and as Dyson and Todd (2010) note, it was labour and resource intensive. This was true for both the evaluation team and the stakeholders, which some might feel diminishes resources for the implementation of the initiative (Cole, 2003). Yet, the benefits of adopting a ToC approach manifest when stakeholders are intensively consulted in its design and have a deeper level of engagement with the intervention itself (De Silva, et al, 2014). We feel this was achieved in this evaluation and the ToC was critical to evidencing the budding success of the IASU.

One drawback of the close collaboration needed for this evaluation might be that the evaluation team influenced the delivery of the ToC itself. Hence, it could be argued that we became participants in the evaluation through our dialogue with stakeholders and the interim findings we reported. Whilst this negates controlling for externalities in the evaluation, as might be preferred by an experimental approach, it allowed us to document contextual conditions and to collect data that best suited the purposes of the evaluation, rather than retrofitting this at the end of the programme.

Conceptualising ToCs visually as two-dimensional figures means that they are necessarily simplified. On the one hand, this can mean they are easier to understand but, on the other, the complexities and interactions between the different moving parts of a complex social intervention are not captured and 'may conceal some of explanatory power of the causal pathways' (Breuer, et al, 2015). The latter can lead to some interdependencies and interactions being overlooked. Alternatively, a complex ToC with many nested components and feedback loops can make the measurement of interim outcomes within the model and subsequent evaluation problematic (De Silva et al, 2014). Accordingly, we chose to include

A theory of change driven approach to evaluating a multi-agency stalking intervention programme

top level strategic interim outcomes for the quantitative analysis of the hypotheses and used qualitative evidence to support other outcomes.

In conclusion, adopting a ToC driven approach to evaluating the MASIP and IASU proved to be advantageous on grounds identified by Breuer et al. (2015). First, because this approach is compatible with any evaluation framework, including the realist inspired EMMIE framework used herein. Second, for complex social interventions it is a robust alternative to randomised control trials, where such trials are difficult to implement. Finally, describing the process by which the ToC was framed, how the evaluation was designed, and pathways to change measured, allows for the judgement of validity of the findings (Dyson & Todd, 2010) and provides a roadmap for practitioners and policy makers to set up future partnerships.

Conclusion

This evaluation of IASU to reduce the harmful effects of stalking indicated good reason to believe that the multi-agency effort was promising with regards to preventing reoffending. However, the mechanisms through which the IASU achieved that remain unclear, as none of the interventions delivered by this partnership were statistically significantly associated with reoffending as an outcome. We believe a longer follow up period and a larger sample size are essential to demonstrate with some confidence the impact of complex interventions delivered by partnerships like the IASU. Although our attempt to test these theoretically generated hypotheses had limited success due to factors beyond the evaluators' control, we believe that a ToC driven approach combined with an EMMIE inspired evaluation is superior as it can contribute towards specifically identifying whether interventions work, how they work, for whom do they work, and under what conditions. At the very least, our attempt demonstrates the importance of data collection and provision for long term evaluation plans, especially for complex social interventions.

A theory of change driven approach to evaluating a multi-agency stalking intervention programme

References

- Atkinson, M., Jones, M., & Lamont, E. (2007). Multi-agency working and its implications for practice. *Reading: CfBT Education Trust*, 28-42.
- Balloch, S. and Taylor, M., 2001. Introduction. In: S. Balloch and M. Taylor, eds. *Partnership working: policy and practice*. Bristol: The Policy Press, 1–14.
- Barnes, M., Matka, E., & Sullivan, H. (2003). Evidence, understanding and complexity: evaluation in non-linear systems. *Evaluation*, 9(3), 265-284.
- Belur, J., Tompson, L., & Jerath, K. (2019). *MASIP evaluation interim report*. UCL Discovery <https://discovery.ucl.ac.uk/id/eprint/10078857/1/MASIP%20evaluation%20interim%20report%20v2.6.pdf>
- Berry, G., Briggs, P., Erol, R., & van Staden, L. (2011). The effectiveness of partnership working in a crime and disorder context: A rapid evidence assessment. *H. Office (Ed.), Research Report*, 52.
- Björklund, K., Häkkänen-Nyholm, H., Sheridan, L., & Roberts, K. (2010). The prevalence of stalking among finnish university students. *Journal of Interpersonal Violence*, 25(4), 684–698. <https://doi.org/10.1177/0886260509334405>
- Blamey, A., & Mackenzie, M. (2007). Theories of change and realistic evaluation: peas in a pod or apples and oranges?. *Evaluation*, 13(4), 439-455.
- Breuer, E., Lee, L., De Silva, M., & Lund, C. (2015). Using theory of change to design and evaluate public health interventions: a systematic review. *Implementation Science*, 11, 1-17.
- Burnett, R., & Appleton, C. (2004). Joined-up services to tackle youth crime: a case-study in England. *British journal of criminology*, 44(1), 34-54.
- Chapman, J., Higgins, A. and Hales, G. (2017). Slough site report. Police Effectiveness in a Changing World. Police Foundation. Available at: <http://www.police-foundation.org.uk/publication/police-effectiveness-changing-world-slough-site-report/>
- Cho, H., Hong, J. S., & Logan, T. K. (2012). An ecological understanding of the risk factors associated with stalking behavior: Implications for social work practice. *Affilia*, 27(4), 381-390.
- Cleaver, K., Maras, P., Oram, C., & McCallum, K. (2019). A review of UK based multi-agency approaches to early intervention in domestic abuse: Lessons to be learnt from existing evaluation studies. *Aggression and violent behavior*, 46, 140-155.
- Cole, M. (2003). The Health Action Zone initiative: lessons from Plymouth. *Local Government Studies*, 29(3), 99–117.

A theory of change driven approach to evaluating a multi-agency stalking intervention programme

- Connell, J. P., & Kubisch, A. C. (1998). Applying a theory of change approach to the evaluation of comprehensive community initiatives: progress, prospects, and problems. *New approaches to evaluating community initiatives*, 2(15-44), 1-16.
- Cox, C. (2019). Policing a diverse society: the community based rationale for multi-agency working. In *Multi-Agency Working in Criminal Justice* (pp. 105-120). Policy Press.
- Crawford, A. (1994). The partnership approach to community crime prevention: corporatism at the local level?. *Social & legal studies*, 3(4), 497-519.
- De Silva, M. J., Breuer, E., Lee, L., Asher, L., Chowdhary, N., Lund, C., & Patel, V. (2014). Theory of change: a theory-driven approach to enhance the Medical Research Council's framework for complex interventions. *Trials*, 15(1), 1-13.
- Dyson, A., & Todd, L. (2010). Dealing with complexity: theory of change evaluation and the full service extended schools initiative. *International Journal of Research & Method in Education*, 33(2), 119-134.
- Firth, D. (1993). Bias reduction of maximum likelihood estimates. *Biometrika*, 80(1), 27-38.
- Funnell, S., and P. Rogers. 2011. *Purposeful programme theory: Effective use of theories of change and logic models*. San Francisco: Jossey-Bass
- Her Majesty's Inspectorate of Constabulary Fire and Rescue Services & Her Majesty's Crown Prosecution Service Inspectorate. (2017). Living in fear—The police and CPS response to harassment and stalking. A joint inspection by HMIC and HMCPSI. London, England: Her Majesty's Inspectorate of Constabulary Fire and Rescue Services. Retrieved from <http://www.justiceinspectors.gov.uk/hmic/publications/living-in-fear-the-police-and-cps-response-to-harassment-and-stalking/>
- Herbert, J. L., & Bromfield, L. (2019). Better together? A review of evidence for multi-disciplinary teams responding to physical and sexual child abuse. *Trauma, Violence, & Abuse*, 20(2), 214-228.
- Hester, M., Eisenstadt, N., Jones, C. and Morgan, K. (2017). Evaluation of the Drive Project - a pilot to address high-risk perpetrators of domestic abuse Year 1 Feasibility Study. University of Bristol.
- Higgins, A., Hales, G., & Chapman, J. (2017). *Police effectiveness in a changing world: Luton site report*. Police Foundation.
- Hobson, J., Lynch, K., Payne, B., & Ellis, L. (2021). Are police-led social crime prevention initiatives effective? A process and outcome evaluation of a UK youth intervention. *International criminal justice review*, 31(3), 325-346.

A theory of change driven approach to evaluating a multi-agency stalking intervention programme

- Holt, T. J., Lee, J. R., Liggett, R., Holt, K. M., & Bossler, A. (2019). Examining perceptions of online harassment among constables in E&W. *International Journal of Cybersecurity Intelligence & Cybercrime*, 2(1), 24-39.
- Jamal, F., Fletcher, A., Shackleton, N., Elbourne, D., Viner, R., & Bonell, C. (2015). The three stages of building and testing mid-level theories in a realist RCT: a theoretical and methodological case-example. *Trials*, 16(1), 1-10.
- Jerath, K., Tompson, L., & Belur, J. (2022a). Risk management in stalking victims: a multi-agency approach to victim advocacy. *Journal of interpersonal violence*, 37(11-12), NP8989-NP9015.
- Jerath, K., Tompson, L., & Belur, J. (2022b). Treating and managing stalking offenders: findings from a multi-agency clinical intervention. *Psychology, Crime & Law*, 1-24.
- Johnson, S. D., Tilley, N., & Bowers, K. J. (2015). Introducing EMMIE: An evidence rating scale to encourage mixed-method crime prevention synthesis reviews. *Journal of Experimental Criminology*, 11(3), 459-473.
- Kelly, R. (2020). The problematic development of the Stalking Protection order. *Modern Law Review*. 83(2), 406-427
- Kemshall, H., & Wood, J. (2013). MAPPA: The management of sex offenders. In *Sex Offenders: Punish, Help, Change or Control?* (pp. 285-304). Routledge.
- Kirkland, J., & Baron, E. (2015). Using a cognitive analytic approach to formulate a complex sexual and violent offender to inform multi-agency working: developing a shared understanding. *Journal of Sexual Aggression*, 21(3), 394-405.
- Korkodeilou, J. (2016). Stalking victims, victims of sexual violence and criminal justice system responses: Is there a difference or just 'business as usual'? *British Journal of Criminology*, 56(2), 256–273. <https://doi.org/10.1093/bjc/azv054>
- Little, M., Kogan, J., Bullock, R., & Van Der Laan, P. (2004). ISSP: An experiment in multi-systemic responses to persistent young offenders known to children's services. *British Journal of Criminology*, 44(2), 225-240.
- MacKenzie, R. D., & James, D. V. (2011). Management and treatment of stalkers: Problems, options, and solutions. *Behavioral Sciences & the Law*, 29(2), 220-239.
- MacKenzie, R. D., McEwan, T. E., Pathé, M. T., James, D. V., Ogloff, J. R. P., & Mullen, P. E. (2009). *Stalking Risk Profile: Guidelines for the assessment and management of stalkers* (1st ed.). Melbourne, Victoria, Australia: StalkInc. and Centre for Forensic Behavioural Science, Monash University

A theory of change driven approach to evaluating a multi-agency stalking intervention programme

Mazerolle, L., Cherney, A., Eggins, E., Hine, L., & Higginson, A. (2021). Multiagency programs with police as a partner for reducing radicalisation to violence. *Campbell Systematic Reviews*, 17(2), e1162.

McKeon, B., McEwan, T. E., & Luebbers, S. (2015). "It's not really stalking if you know the person": Measuring community attitudes that normalize, justify and minimise stalking. *Psychiatry, Psychology and Law*, 22(2), 291-306.

McNamara, C. L., & Marsil, D. F. (2012). The prevalence of stalking among college students: The disparity between researcher- and self-identified victimization. *Journal of American College Health*, 60(2), 168–174. <https://doi.org/10.1080/07448481.2011.584335>

McEwan, T. E., Shea, D. E., Nazarewicz, J., & Senkans, S. (2017). Reassessing the link between stalking and intimate partner abuse. *Partner Abuse*, 8(3), 223-250.

McEwan, T. E., Mullen, P. E., & MacKenzie, R. (2009). A study of the predictors of persistence in stalking situations. *Law and human behavior*, 33, 149-158.

Mullen, P. E., Pathé, M., & Purcell, R. (2000). *Stalkers and their victims*. Cambridge University Press.

Nijdam-Jones, A., Rosenfeld, B., Gerbrandij, J., Quick, E., & Galietta, M. (2018). Psychopathology of stalking offenders: Examining the clinical, demographic, and stalking characteristics of a community-based sample. *Criminal Justice and Behaviour*, 45(5), 712–731. <https://doi.org/10.1177/0093854818760643>

Ostermeyer, B., Friedman, S.H., Sorrentino, R. and Booth, B.D., (2016). Stalking and violence. *Psychiatric Clinics*, 39(4), pp.663-673.

Owens, J. G. (2016). Why Definitions Matter: Stalking Victimization in the United States. *Journal of Interpersonal Violence*, 31(12), 2196 -2226.

Pathé, M. T., Haworth, D. J., Goodwin, T. A., Holman, A. G., Amos, S. J., Winterbourne, P., & Day, L. (2018). Establishing a joint agency response to the threat of lone-actor grievance-fuelled violence. *The Journal of Forensic Psychiatry & Psychology*, 29(1), 37-52.

Pathé, M., Mackenzie, R., & Mullen, P. E. (2004). Stalking by law: Damaging victims and rewarding offenders. *Journal of law and medicine*, 12, 103-111.

POSTnote: Parliamentary Office of Science and Technology (2018). Stalking and Harassment. (POST Publication No. 592). Retrieved from <https://post.parliament.uk/research-briefings/post-pn-0592/>

Pawson, R., & Tilley, N. (1997). An introduction to scientific realist evaluation. *Evaluation for the 21st century: A handbook*, 1997, 405-18.

A theory of change driven approach to evaluating a multi-agency stalking intervention programme

- Reeves, C. (2013). How multi-agency are multi-agency risk assessment committees?. *Probation Journal*, 60(1), 40-55.
- Robinson, A. L. (2004). Domestic Violence MARACs (Multi-Agency Risk Assessment Conferences) for Very High-risk Victims in Cardiff, Wales: A Process and Outcome Evaluation. Cardiff University School of Social Sciences.
- Robinson, A. L., & Clancy, A. (2021). Systematically identifying and prioritising domestic abuse perpetrators for targeted intervention. *Criminology & Criminal Justice*, 21(5), 687-704.
- Rogers, P. (2014). Theory of change. *Methodological briefs: Impact evaluation*, 2(16), 1-14.
- Rosenbaum, D. P. (2002). Evaluating multi-agency anti-crime partnerships: Theory, design, and measurement issues. *Crime prevention studies*, 14, 171-225.
- Sloper, P. (2004). Facilitators and barriers for co-ordinated multi-agency services. *Child: care, health and development*, 30(6), 571-580.
- Smith, P., Sugiura, L., Karagiannopoulos, V., & Howard, S. (2017, August). A multi-agency and collaborative approach to tackling cybercrime. In *IAFS 2017: 21st Triennial Meeting of the International Association of Forensic Sciences 2017*.
- Stein, D., and C. Valters. 2012. *Understanding theory of change in international development*. Justice and Security Research Programme Paper 1. London: London School of Economics and Political Science.
https://eprints.lse.ac.uk/56359/1/JSRP_Paper1_Understanding_theory_of_change_in_international_development_Stein_Valters_2012.pdf
- Storey, J. E., & Hart, S. D. (2011). How do police respond to stalking? An examination of the risk management strategies and tactics used in a specialized anti-stalking law enforcement unit. *Journal of Police and Criminal Psychology*, 26, 128-142.
- Taylor-Dunn, H., Bowen, E., & Gilchrist, E. A. (2018). Reporting harassment and stalking to the police: A qualitative study of victims' experiences. *Journal of Interpersonal Violence*, 36(11-12), NP5965–NP5992. <https://doi.org/10.1177/0886260518811423>
- Tompson, L., Belur, J., & Jerath, K. (2020). *MASIP evaluation final report*. UCL Discovery <https://discovery.ucl.ac.uk/id/eprint/10097009/>
- Tompson, L., Belur, J., & Jerath, K. (2021). A victim-centred cost–benefit analysis of a stalking prevention programme. *Crime Science*, 10, 1-11.
- Vander Laenen, F., & Persac, N. (2014). Whose claim is legitimate anyway? Negotiating power in inter-agency collaboration. *Legitimacy and Trust in Criminal Law, Policy and Justice: Norms, Procedures, Outcomes*, 111-129.

A theory of change driven approach to evaluating a multi-agency stalking intervention programme

Van Tulder, R., & Keen, N. (2018). Capturing collaborative challenges: Designing complexity-sensitive theories of change for cross-sector partnerships. *Journal of Business Ethics*, 150(2), 315-332.

Weiss, C. H. (1997). How can theory-based evaluation make greater headway? *Evaluation review*, 21(4), 501-524.

Wheatley, R. C. (2019). *What drives men who commit stalking offences and how practitioners can best respond to their needs*. Nottingham Trent University (United Kingdom).

Wilkinson, H., D. Hills, A. Penn, and P. Barbrook-Johnson. (2021). Building a system-based theory of change using participatory systems mapping. *Evaluation* 27 (1): 80–101.

Pre-print

A theory of change driven approach to evaluating a multi-agency stalking intervention programme

Appendix

Study	Study country	Sample size	How measured	Re-offending rate	Re-offending type
Eke et al. 2011	North America	78	(Sub-sample of Mohandie et al. 2006). Police contacts after threat assessment and legal intervention.	56% (n=44)	Recidivism
Foellmi et al. 2016	North America	89	Evidence of post-assessment stalking behaviour or harassment after psychological intervention from self-reports and criminal records.	35% (n=31)	Recidivism
Hehemann et al. 2017	The Netherlands	115	Report of stalking made to police within six months of index offence.	40% (n=46)	Persistence
	England	222		44.1% (n=98)	Persistence
James et al. 2010	Australia	140	When a subject re-contacted their target (no ex-intimates in sample). No timeframe in English sample who received a legal warning. Australian sample received a psychological intervention.	50% (n=70)	Persistence over 12 weeks
				27.1% (n=38)	Persistence over 52 weeks

A theory of change driven approach to evaluating a multi-agency stalking intervention programme

Malsch et al. 2011	The Netherlands	709	New convictions for people already convicted of stalking, over 4 years or more.	11% (n missing)	Recidivism (stalking)
				24% (n missing)	Recidivism (stalking related)
McEwan et al. 2009	Australia	200	Prior stalking, to the same or other victim, disclosed during assessment with mental health clinician during a psychological intervention.	33% (n=65)	Recidivism
				9.5% (n=19)	Recurrence
McEwan et al. 2017	Australia	157	Stalking prior to index offence was taken from criminal records and clinician reports.	10% (n=16)	Persistence/ Recurrence (same victim)
				27% (n=43)	Recurrence (different victim)
McEwan et al. 2018	Australia	235	Stalking charge or restraining order (or similar), to the same or other victim, after psychological treatment. Over an average follow up time of 4 years. 8 offenders stalked both.	14.9% (n=35)	Recurrence - same victim
				13.6% (n=32)	Recurrence - new victim

A theory of change driven approach to evaluating a multi-agency stalking intervention programme

				26.4% (n=62)	Recidivism
McEwan et al. 2019	The Netherlands	70	Police reports by the same or other victim, over an average follow up time of 2.63 years from referral to psychological intervention. 6 offenders stalked both.	40% (n=28) 21.4% (n=15)	Recurrence – same victim Recurrence – new victim
				52.9% (n=37)	Recidivism
Mohandie et al. (2006)	North America	730	When a subject re-contacted their target, subsequent to threat assessment and legal intervention.	60% (n=434)	Persistence
Rosenfeld (2003)	North America	189	Evidence of a second arrest or re-contact of target (based on victim or Probation reports or indicated in clinical interview).	49% (n=93)	Recidivism in whole sample
Rosenfeld et al. 2007	North America	29	Police and/or probation reports for 20 months after a psychological intervention.	0% (n=0/14) 26.7% (n=4/15)	Recidivism for treated stalkers Recidivism for treatment dropouts

Shea 2015	Australia	148	Police charges for stalking or indicative stalking post index offence and post psychological intervention. Sample is cases not offenders.	20.3% (n=30)	Recidivism (stalking charge)
				30.4% (n=45)	Recidivism (indicative stalking charge)
Shea et al. 2018	Australia	146	Stalking post index offence and post psychological intervention was taken from criminal records and self-reports.	33.6% (n=49)	Recurrence

Table 4 – Summary of re-offending statistics from academic research, with type of re-offending measure.

N.B. recidivism means re-offending over a non-specified period, persistence refers to an ongoing offence to the same victim within 6 months and persistence refers to a break in stalking behaviour at least six months before resuming (to the same or a different victim).

These definitions are taken from McEwan et al. (2009).

Pre-print