

Measuring the Success of Counter-Trafficking Interventions in the Criminal Justice Sector: Who decides—and how?

Anne T. Gallagher and Rebecca Surtees

Abstract

Global concern about human trafficking has prompted substantial investment in counter-trafficking interventions. That investment, and the human rights imperatives that underpin counter-trafficking work, demand that interventions demonstrate accountability, results and beneficial impact. How this can happen in practice is complicated and contested. This article, which considers success measurements with respect to criminal justice interventions, seeks to cut through the complexities presented by multiple theories and elaborate methodologies by focusing on one key issue: who decides success, and how? A review of evaluation reports and interviews with practitioners confirm that determinations of success (or failure) will vary according to: (i) who one consults and their role in the intervention; (ii) the criteria against which success is measured; and (iii) the assumptions that are built into that criteria. Each aspect is considered with reference to examples and insights drawn from recent practice. A major finding of the article is that the lack of an overarching vision of what “success” might look like allows mediocre or even harmful interventions to flourish and good work to go unrecognised and unrewarded.

Key words: trafficking, human trafficking, criminal justice, monitoring, evaluation, impact assessment

Introduction

Global concern about human trafficking has nurtured great legal and normative change. Over the past decade, it has also prompted substantial investment in counter-trafficking interventions by intergovernmental organisations, states and civil society. Initial waves of intervention took place in a performance evaluation vacuum. This was noted by, amongst others, the United States Government Accountability Office which, in 2006, criticised the absence of ‘measurable goals and associated indicators to evaluate the overall effectiveness of [US] efforts to combat trafficking abroad’ and echoed the State Department’s Inspector-General in calling for ‘performance indicators to compare progress in combating trafficking from year to year’.¹ Time, experience and heightened expectations about what can be achieved have fed a demand for tools and mechanisms to make sense of the problem and validate what is being done by, for example, measuring the true extent of trafficking or evaluating the absolute and comparative worth of an individual state’s response or a particular intervention. A rapid rise in the formulation and application of “success indicators” is one manifestation of the new environment within which counter-trafficking is being discussed and targeted.²

Equally relevant are the recent but increasingly frequent calls for greater transparency and accountability within the counter-trafficking sector, including through rigorous impact evaluation.³ It is not difficult to sustain a strong argument that counter-trafficking interventions, including those in the criminal justice sector, should be carefully monitored and evaluated. Certainly, the human rights imperatives that underpin counter-trafficking work and the significant investment of public resources demand that interventions demonstrate accountability, results and beneficial impact. How this can happen in practice is more complicated, and there has been relatively little analysis of the practical

¹ United States Government Accountability Office (USGAO), *Human Trafficking: Better Data, Strategy, and Reporting Needed to Enhance U.S. Antitrafficking Efforts Abroad*, USGAO, Washington, DC, 2006, p. 25.

² See: A Gallagher and J Chuang, ‘The Use of Indicators to Measure Government Responses to Human Trafficking’ in B Kingsbury and S Merry, *Indicators as a Tool of Global Governance*, Oxford University Press, forthcoming 2012.

³ G Baumann, ‘Measuring Effectiveness of Counter-Slavery Work’, *Free the Slaves*, Washington, DC, 2006; G Danailova-Trainor, ‘Evaluation of Counter-Trafficking Programmes and Research’ in C Aghazarm, F Laczko, et al., *Human Trafficking: New directions for research*, International Organization for Migration (IOM), Geneva, 2008, pp. 44–56; M Dottridge, ‘Research Needs

issues and challenges that may arise in efforts to evaluate counter-trafficking work. In our view, discussions of counter-trafficking evaluations should identify and acknowledge these obstacles and constraints, and consider what might be done to accommodate or overcome them.

Impact and effectiveness evaluation in the context of international development is complicated and contested. Multiple theories and elaborate methodologies abound, and these can present a daunting impediment to those seeking practical guidance in determining “what works”. This article has sought to cut through some of these complexities by focusing on several basic issues that are directly implicated in evaluating counter-trafficking interventions in the criminal justice sector. The key question is: who decides, and how? Determinations of success (or failure) are likely to vary according to who one consults and their role in the intervention, as well as, most importantly, the criteria against which success is measured and the assumptions that are built into that criteria. By addressing this

Concerning the Monitoring, Evaluation and Impact Assessment of Both Research about Human Trafficking and Projects and Programmes to Address Human Trafficking’ in C Aghazarm, F Laczko, et al., *Human Trafficking: New directions for research*, IOM, Geneva, 2008, pp. 57–71; M Dottridge, ‘Measuring Responses to Trafficking in Human Beings in the European Union: An assessment manual’, European Commission (Paper prepared for the Directorate-General Justice, Freedom and Security European Commission), 2007; M Friedman, ‘This story could have been written 10 years ago. We need to do more’ Rights Work: www.rightswork.org, online publication, 15 August 2011; Global Alliance Against Traffic in Women (GAATW), *Collateral Damage: The impact of anti-trafficking measures on human rights around the world*, GAATW, Bangkok, 2007; GAATW, C Hames, F Dewar, and R Napier-Moore, *Feeling Good about Feeling Bad...A global review of evaluation in counter-trafficking initiatives*, GAATW, Bangkok, 2010; J Intili, et al., *An Evaluation Framework for USAID-Funded TIP Prevention and Victim Protection Programs*, Social Transition Team, Office of Democracy, Governance and Social Transition of the United States Agency for International Development (USAID/E&E/DGST), Washington, DC, 2009; IOM, *Handbook on Performance Indicators for Counter-Trafficking Projects*, IOM, Geneva, 2008; A Jordan, ‘More Funding for Impact Assessments?’ Rights Work: www.rightswork.org, online publication, 15 October 2010; R Rosenberg, *Best Practices for Programming to Protect and Assist Victims of Trafficking in Europe and Eurasia*, USAID/E&E/DGST, Washington, DC, 2008; R Surtees, *Re/integration of Trafficked Persons: Developing monitoring and evaluation mechanisms*, KBF, Brussels, and NEXUS Institute, Washington, 2009; USGAO, *Human Trafficking: Monitoring and evaluation of international projects are limited, but experts suggest improvements*, USGAO, Washington, DC, 2007; E Williamson, et al., *Where is the Research on Human Trafficking and the Evaluation of Counter-Trafficking Efforts?*, ICF International, 2008.

question, the authors seek to contribute, in a practical way, to current discussions about how trafficking-related interventions can be effectively evaluated.

Several limitations of focus deserve to be flagged upfront. The first is the article's attention to the criminal justice side of counter-trafficking responses: interventions that are directed primarily at strengthening the investigation, prosecution and/or adjudication of trafficking-related cases, as well as the applicable legal framework. Typical "criminal justice sector" interventions include support for criminal law reform; training of police, prosecutors and judges; direct support for investigations and prosecutions; and institutional reform (for example, capacity building of institutions such as specialist investigation units, prosecutorial offices and courts, and development of procedures and protocols governing investigations and prosecutions). While some of the article's findings may be applicable to different areas of counter-trafficking work such as prevention and victim assistance, others are specific to the criminal justice response. A second limitation relates to the focus on externally supported interventions: those that are funded (and typically also developed, managed and evaluated) by bilateral and multilateral donors. While certain conclusions implicate key actors in international development cooperation, such as donors and their evaluators, the issues raised are ultimately relevant to any counter-trafficking intervention, including those initiated, funded and implemented by national governments. It is also important to note that the analysis is limited to just a few of the relevant issues—a consideration of stakeholders, criteria and selected underlying assumptions. A broader study, of which the present article forms part, will extend this analysis to include matters such as data availability, accessibility and quality; resources; and identification of unintended negative consequences.⁴

The article (and the study of which it forms a part) is based on a review of relevant literature including a selection of evaluation reports of recent externally supported interventions with a significant criminal justice focus. Evaluation reports of interventions that did not focus specifically on criminal justice aspects of the counter-trafficking response were also examined for comparative purposes. While

⁴ 'Assessing the Impact and Effectiveness of Counter-Trafficking Interventions in the Criminal Justice Sector: A discussion paper on issues, obstacles and opportunities', Asia Regional Trafficking in Persons Project, forthcoming 2012.

approximately half the reports utilised are publicly available,⁵ the remainder were obtained and used on the basis of confidentiality. In this regard, it is relevant to note that several major donors working in this area do not release their evaluation reports. Confidential interviews were also conducted with twelve individuals currently working in the counter-trafficking field, with professional backgrounds in criminal justice (law enforcement, prosecution, and the judiciary), international law, human rights, development, and monitoring and evaluation.⁶ An important, supplementary source of information was provided by the authors' own experiences of designing, managing, implementing and evaluating counter-trafficking interventions in different regions including Europe, the former Soviet Union and Asia.

1. Whose success?

All counter-trafficking criminal justice interventions involve multiple stakeholders, each of whom will be impacted differently and may therefore have different (and even conflicting) views on what constitutes “success”. Externally supported interventions increase the range of stakeholders who will have a perspective on whether and why

⁵ P Cunningham and S Hong, United Nations Inter-Agency Project on Human Trafficking in the Greater Mekong Sub-region (UNIAP) Phase III (2007–2010), Mid-Term Evaluation Report Final Draft, March 2009, executive summary available from: http://www.no-trafficking.org/reports_docs/uniap_mte_execsum.pdf, full report on file with the authors; International Labor Office, Independent Evaluation Report: Mid Term Evaluation: ILO Combating Forced Labour and Trafficking of Indonesian Migrant Workers, available from: www.norad.no/en/_attachment/106242/binary/5825?download; A Jones, et al., *An Evaluation of the International Justice Mission's "Project Lantern": Assessment of five-year impact and change in the public justice system*, Philippines, 2010, http://www.ijm.ca/pageid/236#_report; NORAD, *Evaluation of the International Organization for Migration and its Efforts to Combat Human Trafficking*, 2010, <http://www.iom.int/jahia/webdav/shared/shared/mainsite/activities/ct/Evaluation-IOM-CT.pdf>; United Nations Office on Drugs and Crime, *Evaluation of the United Nations Global Initiative to Fight Human Trafficking (UN.GIFT)*, February 2011, http://www.unodc.org/documents/commissions/WG-GOVandFiN/UN.GIFT_Evaluation_-_final_draft_24Feb2011.pdf; United States Department of Labor, *Independent Final Evaluation of the Thailand Sex Trafficking Task Force: Prevention and placement program* (International Justice Mission), 2006, http://www.dol.gov/ilab/projects/asia/Thailand_TraffickingTaskForce_feval.pdf; USAID, *International Justice Mission: Counter-trafficking program in Cambodia: Assessment*, 2006, http://pdf.usaid.gov/pdf_docs/PNADG806.pdf.

⁶ Audio recordings and transcripts of all interviews are on file with the authors and journal editors.

the intervention has been successful or not. In this context therefore, the term “stakeholder” potentially includes the recipient government/agency; implementing partner agencies and officials including investigators, prosecutors and judges; donors; trafficking victims and victim support agencies; suspects and convicted offenders; evaluators; and technical experts. Each of these stakeholders may assess the success (or failure) of a particular criminal justice project or intervention in the counter-trafficking field using different standards and criteria that reflect their own positions, interests and assumptions. For example:

- *Target government/agency*: Has the intervention reflected well on the government/agency—for example, is the government/agency able to claim credit for any positive results? Did the intervention help ameliorate internal or external criticisms directed at the government/agency? Did it improve the capacity of the government/agency to investigate and prosecute trafficking-related crimes or have broader positive effects on capacity? Did it support implementation of national laws and realisation of national plans or policies? Did the intervention result in any unintended and/or negative consequences?
- *Implementing partner criminal justice agencies (e.g. police unit receiving victim interview facilities, prosecutors receiving training)*: Has the intervention raised the profile of the agency in a positive way? Do the results of the intervention reflect well on practitioners? Has there been a measurable change in relation to the criteria by which success is judged internally, such as victim rescues, arrests, prosecutions and convictions? Are perpetrators being deterred from committing future offences? Has there been an increase in the number of identified or assisted victims who are willing to cooperate? Are cases being processed more quickly? Has cross-border cooperation increased in ways that facilitated the work of the agency? Did the intervention support a focus on high-end exploiters or just the arrest and prosecution of small-time offenders? Has the intervention supported common standards and approaches that will encourage greater regional cooperation? Is the donor satisfied with the results and thereby likely to provide further assistance? Did the intervention result in any unintended and/or negative consequences that could reflect badly on the implementing partner?
- *Donor*: Has the intervention secured clear, unambiguous results that can be quantified, measured and reported? Can it be justified as representing value for money? Has the intervention

strengthened important political or strategic relationships (for example, with a partner country or regional institution)? Has the intervention served other political or strategic goals—for example, reducing the flow of trafficked persons into the donor state? Has the intervention reflected well on the donor? Did the intervention result in any unintended and/or negative consequences that could reflect badly on the donor? Can the intervention be replicated by other agencies and/or by the donor in different settings?

- *Victims:* Did the intervention facilitate accurate and timely identification of victims, their escape or removal from exploitation and protection from further harm? Did victims receive the assistance and protection they require? Did the intervention facilitate greater criminal justice awareness of victim rights? If so, did greater awareness of victim rights translate into appropriate treatment of victims? Were victims treated sensitively in the legal process with full respect of their rights? Did the intervention facilitate positive (or at least minimally traumatic) victim involvement in the investigation and prosecution of their exploiters? Was the legal process comprehensible to the victim? Did it support access to entitlements such as compensation and right to stay in the destination country? Did victims value the criminal justice response? Did the intervention result in any unintended and/or negative consequences, for example, increased likelihood of prosecution for status-related offences such as illegal work or illegal stay?
- *Victim support agencies:* Victim support agency views of success may not always align with those of trafficked persons. In addition to the above success criteria, victim support agencies may consider whether the intervention reflected well on them and their work; whether it advanced particular institutional or programmatic goals; whether it brought in additional funding or increased the prospect of future funding; and whether it improved working relations with criminal justice agencies.

If the perspectives of suspects and offenders were considered important to judgments of success, relevant criteria could include whether the intervention supported fair trial rights and proportionate sentencing; whether it contributed to correct procedures being followed in the legal process; and whether the treatment of suspects has improved. Additional and different perspectives may include those of

the evaluators, whose primary focus would likely be on the extent to which the intervention secured its stated aims, performed against its predetermined indicators and demonstrated value for money. Technical experts and others attached to the implementing agency might have a different view of what constitutes success, which may (or may not) intersect and overlap with some of the views outlined above.

Finally, in this area, it is necessary to acknowledge the existence of silent stakeholders who, while not formally associated with the intervention, nevertheless make their own judgment of success and potentially sway the perspectives of others. A conspicuous example is the United States government, which engages in a controversial but highly influential annual evaluation of state responses to trafficking.⁷ While the task of evaluating state responses to trafficking is different from evaluating the impact of a specific intervention, there are important connections and overlaps. The United States mechanism uses a range of success criteria that prioritises stronger criminal justice responses, evidenced by institution building and increases in prosecutions and convictions. This process may well affect the perspective of key stakeholders, such as recipient states and criminal justice agencies, on the success of any external intervention. Its role as a major donor in the area of criminal justice responses to trafficking has provided the United States government with an additional avenue through which to advocate its particular vision of “success”. Australia provides another example of a donor with a strong international presence in this area and the demonstrated capacity to influence how states—and even a regional grouping of states—determine success.⁸

⁸ See: Association of South East Asian Nations (ASEAN), *Progress Report on Criminal Justice Responses to Trafficking in Persons in the ASEAN Region*, ASEAN, Jakarta, 2011 (examining state and regional performance against detailed performance criteria developed with Australian government support).

⁷ In 2000, the United States Congress passed the *Victims of Trafficking and Violence Protection Act* (TVPA) requiring its State Department to issue annual reports describing ‘the nature and extent of severe forms of trafficking in persons’ and assessing governmental efforts across the world to combat such trafficking against criteria established by United States law. The TVPA lays down “minimum standards” for the elimination of trafficking as well as detailed criteria for evaluating the performance of states. The reports use a ranking system to classify all states reviewed into four tiers of counter-trafficking compliance. Any bottom-tier state, being one that does not comply with the minimum standards and that is not making significant efforts to do so, may be subject to a range of economic sanctions. See further, A Gallagher, ‘Improving the Effectiveness of the International Law of Human Trafficking: A vision for the future of the U.S. TIP Reports’ 12 *Human Rights Review*, 2011, http://works.bepress.com/anne_gallagher/16/.

2. Dealing with conflicting and divergent opinions on success

Within stakeholder groups, there may be internal differences as to what constitutes a successful outcome and, even at the individual level, views on what constitutes “success” may change over time and in response to different contexts. Some victims, for example, may view the conviction of their exploiter after a long and personally difficult legal process, as a positive (and even empowering) result. For others, the process may be so profoundly disorienting and disempowering that even a successful prosecution cannot assuage their dissatisfaction. As noted above, victim support agency views of success may not align with those of trafficking victims. The issue of shelter detention for victims of trafficking provides a relevant example, with a recent study confirming a sharp divergence in views amongst support agencies, and between support agencies and victims regarding the value and impact of shelter detention.⁹

The perspectives of certain stakeholders are often prioritised in evaluating criminal justice interventions in the counter-trafficking sector, while other perspectives are marginalised or discounted. In many externally funded interventions, donors exercise a tight grip over the structure, composition and implementation of evaluations, not least to ensure that their views and interests are given due attention. In other interventions, it is the implementing agency that controls the evaluation process—a conflict of interest that is rarely remarked upon. For example, the evaluation of the UN Global Initiative to Fight Trafficking (UN GIFT) was organised by and involved UNODC, the agency responsible for implementing this large, multi-donor programme. Donor or implementing agency control over the evaluation process can translate into a capacity to ensure that inconvenient evidence-based conclusions are downplayed or ignored.¹⁰

⁹ A Gallagher and E Pearson, ‘The High Cost of Freedom: A legal and policy analysis of shelter detention for victims of trafficking’ *Human Rights Quarterly*, vol. 32, 2010, pp. 73–114.

¹⁰ M Dottridge, ‘Research Needs Concerning the Monitoring, Evaluation and Impact Assessment of Both Research about Human Trafficking and Projects and Programmes to Address Human Trafficking’ in C Aghazarm, F Laczko, et al., *Human Trafficking: New directions for research*, IOM, Geneva, 2008, pp. 57–71, 65–66.

Research for the present article found that it is the views of the most direct and easily accessible programme beneficiaries (for example, the criminal justice officials and agencies that are the target of the intervention) that are most consistently sought out and taken into account. This can skew evaluation outcomes, particularly when the claims of such groups are not tested against the views of others, as these programme beneficiaries may be reluctant to criticise an intervention because of appreciation for assistance provided and/or a fear that such assistance could be withdrawn if negatively assessed.¹¹ In none of the examined criminal justice evaluations were the views of victims or suspects sought or considered. Of course, a failure to consult these groups does not automatically mean that such perspectives are completely ignored. For example, most of the criminal justice experts interviewed for this study explained that their assessment of the success or failure of a project, activity or single case was tied to their perception of how the intervention impacted victims and/or suspects. It is essential to acknowledge, however, that this perception may differ substantially from how victims or suspects themselves assess success or failure.

Cost, time and access considerations clearly play a part in deciding who, among stakeholders, should be involved in the evaluation process. Access to some stakeholders, such as criminal justice practitioners and victim support agencies, may be easier, less costly and less time consuming to secure than others. Certainly, the involvement of trafficked persons and suspects in any evaluation is likely to be very complex and expensive. For example, sufficient time is needed to identify a representative sample of the target group, secure informed consent and develop rapport necessary for safe, ethical and useful interviewing. Such factors may go some way towards explaining why these more complicated stakeholders are routinely omitted from consultations.¹² The possibility that the more difficult-to-reach groups are accorded a lower priority, and that resource and other constraints are used as an excuse to exclude them, should also not be discounted.¹³

¹¹ A Brunovskis and R Surtees, *Leaving the Past Behind? When trafficking victims decline assistance*, NEXUS Institute, Oslo, Fafo and Vienna, 2007, p. 105.

¹² These practical obstacles are rarely acknowledged and it is ethical issues, such as the risk of re-traumatisation, that are most commonly cited as a reason not to interview victims directly. While certainly a critical consideration in any engagement with trafficked persons, it is important that ethical concerns do not become an excuse for avoiding work that is sensitive and difficult. Trafficked persons have a right to be part of any evaluation, and the risk of wasteful and even harmful interventions is clearly increased when the experiences, assessment and needs of victims are not part of an evaluation.

Results of an intervention can be interpreted differently, depending on individual perspective, background and expertise, with widely diverging assessments of what constitutes a “successful” outcome. Consider the real-life example of an externally supported cross-border law enforcement cooperation mechanism that, somewhat unexpectedly, was linked to a significant number of victim rescues. While the donor regarded victim rescues to be a strong indicator of success, technical experts within the project had a different, more nuanced view. They pointed out that this figure revealed nothing about: the number of victims who were not rescued (including those who continued to be exploited with the knowledge of national counterparts with whom the intervention was working); the extent to which “rescues” translated into victims being offered and accepting assistance; the cost of the cooperation process relative to alternative strategies; the time and effort put into making it work; and the apparent inability of the cooperation process to facilitate the identification and arrest of suspects.

Another example of conflicting perspectives of success is provided by the common measure of trafficking prosecutions, considered further in the following section. For present purposes, it is relevant to note that the arrest, prosecution and conviction of traffickers are heavily dependent on the cooperation of victims. However, trafficking victims’ involvement in the criminal justice process may involve significant and on-going risks to their personal safety and physical and emotional well-being, for little or no personal benefit. As a result, victim cooperation must either be compelled or secured very carefully, through, in the words of an experienced law enforcement practitioner, ‘a mixture of encouragement, persuasion and lack of information’. Under these circumstances, it is very likely that trafficked persons and criminal justice agencies will feel very differently about the impact and value of a criminal justice intervention that seeks to enhance the capacity of investigators and prosecutors to bring victims into the criminal justice process as witnesses. That conflicting perspective is likely to be lost, at least in the context of formal performance evaluations conducted by criminal justice agencies, (i) because of the routine failure to seek the views of victim-witnesses; and (ii) because a primary goal (increased victim involvement in prosecutions) was likely set with little consideration of victim perspective.

¹³ M Bamberger and J Rugh, ‘Real World Evaluation. Working under budget, time, data and political constraints’, American Evaluation Association: Professional Development Workshop Session 21, 2008, 7.

One interviewee characterised the current situation, with its multiple stakeholders and multiple perspectives, as creating ‘a messy soup of expectations’ that is extremely difficult, if not impossible, to manage successfully. Certainly, a strong evaluation should take into account different views and positions. However, not all stakeholders are equal and not all perspectives are, or should be, of equal weight. For example, international human rights law provides a strong framework within which many aspects of a counter-trafficking intervention could and should be measured.¹⁴ An approach to evaluation that prioritises human rights of victims and the administration of criminal justice can provide the framework within which different considerations of success can be weighed. As a practical matter, an appropriate balancing of interests also requires that criteria for success (discussed further below) are transparent and logical; that all major stakeholders are identified and consulted; and that different perceptions of the programme and its success are openly acknowledged in the evaluation, even when this diverges from the general assessment.

3. Deciding the criteria for success

In some cases examined, criteria for success of a counter-trafficking criminal justice intervention were not articulated at all. However, such criteria are typically made formal and explicit, attached to the relevant project or programme as objectives, together with predetermined indicators. Of course, this does not prevent unarticulated criteria from influencing an evaluation. For example, while pre-determined indicators are unlikely to extend to the question of whether the intervention has served a donor’s political or strategic interests, this measure of success may nevertheless be deeply entrenched in the evaluation process and its conclusions. Nevertheless, it is the intervention’s formal indicators that, at least in principle, determine whether or not “success” has been achieved.

Several experts interviewed for this study identified the absence of an agreed set of standards that define a successful counter-trafficking criminal justice response as a major problem—an issue that is raised in the conclusion of this article. One participant, for example, noted

¹⁴ See below note 30 and accompanying text.

the lack of agreement on what constitutes progress: ‘It’s not that we can’t measure impact because of our starting point, it’s because we don’t actually know our end point, where we’re trying to get to.’ Other problems include the adoption of unrealistic or unclear objectives, formulated in ways that make evaluation of outcomes or impact virtually impossible. Examples cited included high-sounding objectives (for example, “to eliminate trafficking”) or those that are extremely vague (for example, “to develop more effective criminal justice responses”). Interviewees also pointed to an apparently widespread preference for weak or easily attainable goals—such as whether the programme has been implemented as planned—rather than whether it is able to demonstrate change and impact.¹⁵ This preference for process-related goals over those focused on outcomes enables stakeholders, most particularly implementing agencies and donors, to retain control over the “success story” in a highly uncertain environment by ensuring that evaluative measurement extends only to aspects that are under the direct control of the project or intervention.

Definitions of success in criminal justice responses to trafficking almost invariably fall back on what can be measured, hence the strong preference, particularly among donors, for quantifiable indicators such as number of arrests or prosecutions, number of officials trained, instances of cross-border investigation cooperation, number of special interview suites established, number of relevant treaties ratified, etc. One criminal justice specialist expressed frustration that ‘results have become so important—and the more concrete they are, the more they are appreciated’, even when the basis of these results may be weak or flawed. Another highlighted how this preference for the measurable can have negative implications in an environment where ‘even bad data is [considered] better than no data at all’.¹⁶ One evaluation expert

¹⁵ Rosenberg notes that this is also the case in almost all of the protection programmes reviewed for an assessment of trafficking programmes in the Europe and Eurasia region: R Rosenberg, *Best Practices for Programming to Protect and Assist Victims of Trafficking in Europe and Eurasia*, USAID/E&E/DGST, Washington, DC, 2008, p. 34.

¹⁶ See, for example, the response of a major United Nations agency to data quality problems in the context of its global trafficking report: ‘A poor indicator is better than no indicator as long as it is not represented as more than it is. Over time, the collection of information from so many different perspectives can, in aggregate, make up for many of the deficiencies of the data itself. Our global data set, reviewed time and again, can indeed tell us something more about the trends and patterns of the problem. This information is vital so that, in a world of limited resources, efforts can be focused for maximum effect.’ United Nations Office on Drugs and Crime (UNODC), *Global Report on Trafficking in Persons*, Vienna, 2009, p. 13.

with experience in counter-trafficking interventions criticised the strong donor preference for (often poor quality) quantification over a more sophisticated and robust mix of qualitative and quantitative methods.

4. Interrogating underlying values and assumptions

Views on an intervention's success often depend on underlying values and assumptions. Criminal justice interventions in the counter-trafficking field are rife with untested assumptions about the value of a particular approach, and about stakeholders and beneficiaries that are rarely explicitly articulated. Any credible evaluation should extend to considering whether the underlying assumptions of an intervention are valid.¹⁷ Some of these assumptions, as they relate to how an intervention's success is measured, are considered further below.

Assumption 1: Increases in trafficking-related prosecutions/convictions is a strong and reliable indicator of success

As noted previously, an increase in trafficking-related prosecutions is widely viewed to indicate a more effective criminal justice response to trafficking.¹⁸ Certainly this indicator should not be dismissed outright. The failure of states to arrest and prosecute exploiters has likely contributed to the high levels of impunity currently enjoyed by traffickers and to the widespread denial of justice to victims.

However, an increase in measurable criminal justice activity (arrests, prosecutions and convictions) is a crude and potentially misleading success indicator. For example, changes in prosecution rates may be attributable to new trafficking laws, with offences previously charged under other laws, such as sexual assault, pimping, immigration fraud,

¹⁷ R Surtees, *Re/integration of Trafficked Persons: Developing monitoring and evaluation mechanisms*, KBF, Brussels, and NEXUS Institute, Washington, 2009, p. 19.

¹⁸ United States Department of State, *Trafficking in Persons Reports: 2001-2010*, United States Department of State Washington, DC, (2001)–(2010); UNODC, *Global Report on Trafficking in Persons*, UNODC, Vienna, 2009; IOM, *Handbook on Performance Indicators for Counter-Trafficking Projects*, IOM, Geneva, 2008.

now prosecuted as trafficking. They may also reflect changes in trafficking prevalence as well as the way data is collected. Most importantly, numerical information says nothing about the quality of prosecutions and convictions, discussed further below. A number of respondents argued that the emphasis on increased prosecutions and convictions has resulted in unintended consequences—for example, a focus on “easy” cases such as those involving small-time recruiters, and a misidentification of people smugglers or marriage brokers as traffickers in order to boost the number of trafficking cases pursued.¹⁹

There is some recent acknowledgment of the need to focus on the quality of prosecutions and convictions. Success indicators recently developed through one major criminal justice intervention in South-East Asia and subsequently adopted at the regional level, consider whether the procedural guarantees for a fair trial were provided; whether evidentiary requirements were met; and whether convictions result in adequate and proportionate penalties.²⁰ The International Organization for Migration’s manual of performance indicators suggests convictions as an indicator of success, with the caveat that this assumes a definition of trafficking that is in line with international standards and, critically, a credible justice system that respects due process.²¹

Assumption 2: Declines in prevalence of trafficking indicates positive change

While the prevalence of crime (the extent or proportion of cases in any given population) is not necessarily a routine indicator for criminal justice interventions,²² there appears to be a growing assumption that a decline in the prevalence of trafficking (presumably evidenced by the number of identified victims) is a useful indicator of success. The United States *Trafficking in Persons Report*, for example, implies a

¹⁹ See, for example, Association of South East Asian Nations (ASEAN), *Progress Report on Criminal Justice Responses to Trafficking in Persons in the ASEAN Region*, ASEAN, Jakarta, 2011.

²⁰ *Ibid.*, Chapters 1 and 6.

²¹ IOM, *Handbook on Performance Indicators for Counter-trafficking Projects*, IOM, Geneva, 2008, p. 38.

²² Governance and Social Development Resource Centre (GSDRC), *Helpdesk Research Report: Prevalence as an indicator for counter-forced labour and counter-sex trafficking projects*, GSDRC, 2008, p. 2.

prevalence measurement in establishing, as success criteria, ‘whether [the government under assessment] achieves appreciable progress in eliminating severe forms of trafficking when compared to...the previous year’.²³

The use of prevalence as an indicator is problematic on a number of levels. It does not take into account well-known obstacles and pitfalls associated with quantifying the extent of the trafficking problem.²⁴ It also fails to acknowledge that even if prevalence *could* be accurately measured, changes can easily be attributed to a range of factors beyond the rate at which trafficking is actually occurring. For example, increased attention given to trafficking could result in higher levels of prevalence being recorded. Similarly, increased numbers of identified and assisted victims could well be a partial measure of a country’s efforts to tackle trafficking.²⁵

Further, decreased prevalence may not necessarily signal success but rather the adaptation of trafficking systems to counter-trafficking responses. Consider a recent attempt to use prevalence of child prostitution to measure the success of a criminal justice counter-trafficking intervention targeting child trafficking. The research sought to map prevalence but, as noted by its authors, a measurable change in prevalence may be due to other factors, such as law enforcement pressures that resulted in victims being shifted to new or less visible locations.²⁶

²³ United States Department of State, *Trafficking in Persons Report: 2010*, United States Department of State, Washington, DC, 2011, p. 405.

²⁴ See, for example, USGAO, *Human Trafficking: Better data, strategy, and reporting needed to enhance U.S. antitrafficking efforts abroad*, USGAO, Washington, DC, 2006; E Savona and S Stefanizzi (eds), *Measuring Human Trafficking Complexities and Pitfalls*, Springer Science, New York, 2007; D Feingold, ‘Trafficking in Numbers: The social construction of human trafficking data’ in P Andreas and L Greenhill (eds), *Sex, Drugs and Body Counts: The politics of numbers in global crime and conflict*, Cornell University Press, Ithaca, 2010, p. 46–74; H Ali, ‘Data Collection on Victims of Human Trafficking: An analysis of various sources’ *Journal of Human Security*, vol. 6, no. 1, 2010, p. 55–69.

²⁵ R Surtees, *Second Annual Report on Victims of Trafficking in South-Eastern Europe*, IOM, Geneva, 2005, p. 25.

²⁶ A Jones et al., *An Evaluation of the International Justice Mission’s “Project Lantern”: Assessment of five-year impact and change in the public justice system*, International Justice Mission, 2010, p. 6, 31.

The assumption about prevalence also works in reverse: that increases in prevalence are indicative of failure. However, as one study of impact evaluation has noted, success can also mean doing *less badly*.²⁷ It is not unreasonable to assume that, at least in some parts of the world, the rate at which individuals are being trafficked is increasing. Under such circumstances, the best that an intervention may be able to hope for is to slow down that rate.

Assumption 3: Impact and key success indicators (such as attitudinal change) can be measured accurately and within the life of an intervention

Assumptions and expectations about what can be measured (and when) are often unrealistic. Calls for interventions to demonstrate impact²⁸ often underestimate, or indeed completely overlook, many of the complexities involved, at least within the specialist area of criminal justice responses.

Data quality and accessibility: If evaluation design and implementation are compromised by missing data or poor analysis, the findings will be inadequate and likely irrelevant.²⁹ In many countries, vital information (for example, about rates of trafficking, number, type, quality of investigations, arrests, prosecutions, convictions, protection orders, etc.) is scarce, unreliable and not always verifiable. This severely compromises the capacity of criminal justice interventions to collect baseline information against which future change can be assessed.

²⁷ Center for Global Development, *When Will We Ever Learn? Improving Lives Through Impact Evaluation*, Center for Global Development, Washington, DC, 2010, p. 20.

²⁸ M Friedman, 'This story could have been written 10 years ago. We need to do more.' Rights Work: www.rightswork.org, online publication, 15 August 2011; A Jordan, 'More funding for impact assessments?' Rights Work: www.rightswork.org, online publication, 15 October 2010; GAATW, C Hames, F Dewar, and R Napier-Moore, *Feeling Good about Feeling Bad... A global review of evaluation in counter-trafficking initiatives*, GAATW, Bangkok, 2010.

²⁹ J Intili, et al., *An Evaluation Framework for USAID-Funded TIP Prevention and Victim Protection Programs*, Social Transition Team, Office of Democracy, Governance and Social Transition of the United States Agency for International Development (USAID/E&E/DGST), Washington, DC, 2009, p. 20.

Timing: Several criminal justice professionals interviewed questioned the assumption that impact could be measured, at least within the life of a project or immediately afterwards. They noted that trafficking cases can take years from reception of a complaint to case resolution in court. The results and impacts of interventions that seek to influence criminal justice responses can only fully be measured once the entire process has been completed and in respect to a substantial group of cases. When project objectives are cast, in the words of one evaluation expert, ‘at the upper end of the logic model’, impact becomes ‘virtually impossible to measure unless you put some sort of longitudinal study in place which lasts for a couple of decades’.

Measuring change in attitude and behaviour: Many externally supported criminal justice interventions seek to change the attitudes and behaviour of criminal justice personnel as well as its organisational culture. Such changes do not happen quickly. Any evidence to demonstrate change will likely only come from the way cases are investigated, prosecuted and adjudicated over an extended time period. In short, the true *impact* of capacity building interventions such as training is unlikely to be immediately discernible. One criminal justice adviser highlighted this obstacle, noting that changes in behaviour attributable to skills training in which he was involved were not detectable in counterparts or external observers until at least several years after such training commenced.

Other practical obstacles to measuring changes within the confines of a typical project should not be underestimated. The difficulties are neatly illustrated by one project that has invested heavily in following up law enforcement trainees to assess change in workplace behaviour and thereby move beyond reliance on the one-sided and frequently inconclusive feedback received from trainees themselves. Project staff involved in this evaluation experienced a range of practical difficulties in monitoring workplace performance. Supervisors were rarely available for consultation and, regardless, generally lacked the knowledge and skills to be able to adequately assess the performance and capacities of their subordinates on the relevant issues. In some instances, trained officers had conducted few, if any, trafficking investigations since undertaking the training, thereby preventing a comparison with past performance or behaviour. Compounding these problems was the fact that the basic data required to measure changes in performance or behaviour was often inaccessible or unavailable in a form that could be analysed.

Towards the Future



Recent calls for more rigour and transparency in evaluating the success or failure of counter-trafficking interventions are both welcome and overdue. The need for quality evaluations, including impact evaluations, felt across the development spectrum, appears to be especially acute in the area of criminal justice responses to trafficking. While human exploitation is an age-old phenomenon, this issue has only recently been identified as a priority for states and the international community. States have been developing and adapting their criminal justice responses on the run, often under strong political pressure, and principally through trial and error. Donors and implementing agencies have been operating in a similarly reactive way. This increases the risk of bad decisions and negative consequences. It also makes apparent the urgency of ensuring that robust systems for evaluating interventions are in place and functioning.

A review of the relevant literature confirms that it is much easier to call for more quality in evaluations than it is to actually craft or deliver concrete solutions. Discussions about evaluating counter-trafficking interventions need to move from the important first step of criticism and complaint about what is not being done well (or at all) to serious, informed engagement with obstacles and constraints, and how these can be addressed. A different approach also requires commitment to new levels of openness and transparency. Donors and implementing agencies in particular have a responsibility to contribute to an environment of learning and self-reflection that will help ensure good practices are replicated and poor ones discarded. The publication and dissemination of their evaluation reports should be seen as a critical first step to any serious engagement. It will also be important to recognise that, at least in development terms, many of the challenges of evaluating counter-trafficking interventions are not especially unique and there is considerable space for those working in this area to learn from the experiences and insights of other fields of work.

One of the main impediments to effective evaluations lies in the failure of states, the international community and the counter-trafficking sector to specify and communicate a common vision of what constitutes an effective criminal justice response to trafficking. For example, it is only rarely that success indicators will link to the increasingly sophisticated international legal and policy framework around trafficking that provides detailed and specific guidance on matters such as victims' right to justice and state obligations to exercise due diligence in investigating trafficking, prosecuting perpetrators and

protecting those who have been trafficked.³⁰ The absence of a unified vision, grounded in accepted international rules, manifests itself in the lack of clearly defined end-points that are typical of most criminal justice interventions. It also exacerbates complications such as the presence of multiple and divergent stakeholders, and difficulties in accessing reliable and verifiable data. Even more worryingly, the lack of an overarching vision allows mediocre or even harmful interventions to flourish and good work to go unrecognised and unrewarded. It is a collective responsibility to work towards articulation of this vision and to ensure accountability of those who are provided the resources to promote its realisation.

³⁰ For further explanation of this framework, see A Gallagher, *The International Law of Human Trafficking* (2010), especially at Chapter 7. One example of an explicit attempt to integrate international principles and standards into success criteria for counter-trafficking interventions in the criminal justice sector is provided by ASEAN (2011), *supra* note 10. See also AIM for Human Rights (2010), *The Right Guide: A tool to measure the impact of counter-trafficking laws and policies which, while not focusing specifically on either external interventions or the criminal justice response, provides an important insight into how evaluation can be modified by the integration of a human rights perspective.*



Anne T. Gallagher AO, is Technical Director of the Asia Regional Trafficking in Persons project and an independent scholar and legal adviser. Email: anne.therese.gallagher@gmail.com

Rebecca Surtees is Senior Researcher at NEXUS Institute in Washington, DC.

The views expressed in this article are the authors' and should not be taken to reflect those of the organisations with whom they are or have been associated.