



**A4ID**



**TOOLBOX: MONITORING &  
EVALUATION METHODS  
FOR TRAININGS**



## ABOUT ADVOCATES FOR INTERNATIONAL DEVELOPMENT

Advocates for International Development (A4ID), founded in 2006, is a global charity that believes the law can and should, be used more effectively to advance fair and sustainable development. A4ID aims to inspire and enable lawyers to join the global fight to eradicate poverty by advancing the UN Sustainable Development Goals (SDGs). Through A4ID, the world's top lawyers provide free legal support to organisations, working to advance human dignity, equality and justice. Its work has so far impacted in over 130 countries.

A4ID's Rule of Law Expertise (ROLE UK) Programme is funded by the Foreign, Commonwealth and Development Office. It supports partnerships to provide pro bono legal and judicial expertise with the aim to strengthen the rule of law in official development assistance (ODA)-eligible countries. The Programme's Knowledge Hub provides the pro bono legal sector with access to targeted and relevant information to inform and improve their technical assistance in development contexts.

[www.roleuk.org.uk](http://www.roleuk.org.uk) | [roleuk@a4id.org](mailto:roleuk@a4id.org)

---

## ACKNOWLEDGEMENTS

We would like to acknowledge Naomi Cantor, Learning and Development Officer and Joely Thomas, Monitoring and Evaluation Officer for

their work in developing this toolbox and creating the workshop on which the toolbox is based.

Thanks are also due to Neha Bhansali for layout and design.

---

# CONTENTS

---

**01** Foreword

---

**02** Part 1: Overview of  
Monitoring & Evaluation

---

**12** Part 2: Data Collection  
Methods

---

---

**34** Part 3:  
Evaluate & Adapt

---

**39** Conclusion

---

**40** Further Resources

---

# FOREWORD

A4ID's ROLE UK Programme focuses on developing and strengthening strategic peer-to-peer partnerships to deliver pro bono legal technical assistance in ODA-eligible countries, as well as strengthening practitioners' awareness and skills to support working in the sphere of rule of law and development.

Trainings form a large part of the work undertaken by partner organisations to build capacities in the rule of law sector. Trainings involve the transfer and sharing of knowledge and skills for key practitioners that can support change to institutional practices and advocate wider change in society. Monitoring and Evaluation (M&E) practices, planned well and built into a project's activities, can support the effectiveness of trainings in facilitating this transformation.

Good M&E insights ensure that decisions made in the planning and design of subsequent training activities are based on robust evidence of what has worked well and what has not.

Feedback received by A4ID has shown that practitioners are increasingly recognising the value of deploying robust M&E for assuring the quality of their work and evidencing their achievements. As part of its goal to share and strengthen development thinking among legal and judicial stakeholders, A4ID has created this toolbox to introduce M&E tools and methods that are particularly relevant to trainings, the most commonly used capacity-building method. The tools in the toolbox will support practitioners in creating an M&E strategy that is appropriate, efficient, and effective.

**YASMIN BATLIWALA**

**Chief Executive**



# PART 1: OVERVIEW OF MONITORING & EVALUATION

## INTRODUCTION

A4ID's ROLE UK Programme has produced this toolbox to support those delivering trainings on rule of law issues as part of their international pro bono practice. This resource complements the tools in two earlier papers: '**Maximising Technical Assistance to Improve the Rule of Law**'<sup>i</sup> and the '**Monitoring & Evaluation Toolbox**'<sup>ii</sup>. It integrates the project planning frameworks and methods conveyed within the aforementioned guides and builds upon these. The specificity of this toolbox is that it focuses on M&E for training activities.

The toolbox addresses each stage of enacting an M&E strategy for trainings: planning and preparation, data collection and analysis, and, lastly, the use of findings to adapt programming. It will explore how to make appropriate decisions about M&E approaches, helping trainers to collect and evaluate information about the training and gain insights in an effective and efficient way.

This toolbox is designed to complement the

'Monitoring & Evaluation Methods for Trainings' workshop delivered by A4ID, and most of the tools and approaches covered are introduced in this workshop. It can also be used independently.

## WHO THIS TOOLBOX IS FOR

The toolbox is targeted at legal and judicial specialists undertaking, or seeking to undertake, international legal technical assistance.

The toolbox will be of most relevance to professionals in the legal sector (e.g. lawyers, judges, legal academics). It is especially useful to those who are trying to address specific rule of law issues through targeted interventions, including trainings, and wish to understand and increase the impact of this work.

## WHAT M&E IS AND WHY IT IS SO IMPORTANT

M&E refers to the methods of data collection and analysis that help to evidence and understand the value of a development intervention and the processes of change brought about by it. Once a training has been delivered, it is important to

be assured by robust evidence that the training was successful in so far that it did in fact produce the results expected, both in terms of capacity-building among the participants and having the

wider impact upon the rule of law. This evidence collected also enables practitioners to make informed decisions on how to adapt future projects in order to improve their effectiveness.

**It is important to be assured by robust evidence that the training was successful in so far that it did in fact produce the results expected, both in terms of capacity-building among the participants and having the wider impact upon the rule of law.**





## IMPORTANT PRINCIPLES OF M&E

M&E approaches require that practitioners collect and record information about trainings. This is referred to as data. When considering what data is needed, it is worth remembering that data will only ever provide a limited snapshot of reality. Taking into account the following principles will help to ensure it is the best possible snapshot for informing the analysis of the shorter and longer-term results of training.

Good M&E analysis is dependent upon collecting high quality and sufficient data, which should have the following characteristics:

**Relevant:** For M&E to be effective, the data collected should capture what needs to be known, to make an informed and evidenced-based assessment of the training.

**Reliable:** There should be good grounds to believe that the data collected is accurate. This means that it is not too partial or clearly biased. Collecting data from a variety of sources and including a good-sized sample of responses will enhance validity and reliability. Ideally, comparison of data across different groups or sets will increase confidence that insights are reliable.

**Inclusive:** Data collection methods should take into account any inequalities that may exist within the groups data is collected from. Efforts must be made to ensure that opportunities are provided for all voices to be heard, to ensure that the results reflect different perspectives and experiences.

**Ethical:** There are ethical issues to be considered with data collection methods and how the results

will be used. Data should be collected with informed consent of respondents, and with a right to refuse. Participants should be made aware of how data will be used, and, where possible, data should be recorded anonymously.

**Proportionate:** An M&E plan should be proportionate to the scale of the training. M&E tasks should not be overly burdensome for the facilitators or respondents. Where possible, these activities should be built into the training itself, and harness the easiest opportunities for data collection, bearing in mind relevance and reliability.

**Replicable:** Where possible, it is useful to create tools for data collection which can be easily adapted for future trainings. The results of M&E may then be directly compared across different trainings, strengthening understanding of progress or problems over time. However, M&E tools will need to be reviewed and possibly adapted over time, as the trainings are adapted and improved.

## PLANNING FOR M&E

To ensure that M&E is integrated at each stage of project development and delivery, it should be incorporated into the general planning of the training. This will ensure that the easiest and most effective opportunities to collect data, such as at the training itself, are harnessed, and tools can be easily created and integrated at such opportunities. Formulating a clear M&E plan that incorporates M&E principles should provide a clear roadmap for collecting and analysing data effectively.

## WHAT TO MONITOR AND EVALUATE

An M&E plan is relevant when it ensures that the data to be collected is likely to provide valid and useful insights as to whether a training activity has been effective in achieving the intended short- and long-term results. A training can be deemed as effective in the short-term if the training was delivered as planned and achieved the intended learning outcomes.

When planning the rationale for delivering a training, the long-term anticipated results will require that the learning outcomes are achieved, and that these have translated into a sustained behaviour change in participants. A behaviour change may occur in professional practices as the result of new *capacities* gained at the training. This can also be the result of increasing *motivations* for new forms of action, or increasing

understanding of, or creating new *opportunities* for action. Consideration of these three elements of behaviour change will underpin the rationale for the training activity and the particulars of its design. They are needed to understand whether the training has resulted in the desired behaviour change among participants.

Ideally, a sustained change in the behaviour of participants and potentially those they influence, could translate to long-lasting structural change to the systems and processes for progressing the rule of law. It is therefore important to monitor the long-term impacts of the training or project as a whole, and to evaluate whether any changes that have occurred can be attributed directly to a particular training activity and whether these changes are likely to be sustained over time.

To build an effective M&E plan, the intended short- and long-term results of the project, and how they are expected to be achieved through the training activity, should be clearly articulated. It may be useful to construct what is known as a ‘theory of change’ for your project (see **Monitoring & Evaluation Toolbox**)<sup>iii</sup>.

Additionally, the relevance of the M&E plan may also be guided by the information or results that need reporting to external stakeholders, such as donor organisations. However, it is important that the M&E plan is not only led by these results. A

good M&E plan may also provide relevant insights about *how* and *why* things happened. These kinds of insights will often be drawn from qualitative data, and require investigative follow-ups to initial rounds of data collection. This can help to inform ways to address any barriers or shortcomings that have been identified.

## QUESTIONS TO GUIDE M&E PLANNING AND TOOLS

### THE 3 UNIVERSAL M&E QUESTIONS

There are three core or ‘universal’ questions that are useful to guide the design of an M&E plan for trainings. These questions form the starting point for creating a relevant M&E plan. These questions are:

1. Did we do what we said we would do?
2. Are we creating the change we intended to make?
3. Are these the right things to do?

During the analysis stage, the answers to these questions will inform adaptations to future trainings or to the project as a whole.

### FOCUSED M&E QUESTIONS

Each universal question prompts the generation of a set of sub-questions that focus on the specific elements of the training activity to evaluate. These focused questions will guide the choice of the appropriate methods/tools, the design of the tools themselves and the types of data that need collecting.





## EXAMPLES OF FOCUSED QUESTIONS

### **UNIVERSAL Q.1: Did we do what we said we would do?**

This question regards what happened during the training itself and different aspects of its delivery.

Focused questions should inquire about the use of resources (money, time and expertise), what took place, and the immediate results of the training. These focused questions may also reflect on the quality of different elements of the training.

### **EXAMPLES:**

- Did X number of judges attend the training at each of the sessions?
- Was there an appropriate gender balance in the participant group?
- How much time (including pro bono hours and paid staff time from partnership organisations) was used to deliver the training programme?
- Were the training costs in line with the allocated budget?
- Were the training methods effective?
- Was the trainer's expertise relevant to the context?



## **UNIVERSAL Q.2: Are we creating the change we intended to make?**

This question regards not only the short-term effects of the training, but also the long-term effects, or 'impact'. This will likely concern any behavioural changes of participants that can be attributed to the training, whether learning outcomes were achieved during the training, and whether these had any effect on participants' behaviour. These changes may only come to light some time after the training has happened.

Focused questions about Q.2. should consider short- and long-term changes to behaviours.

For example, after a training, judges could have improved their case management practices. This, in turn, may lead to a change in circumstances where the quantity of cases processed by the courts increases.

### **FOCUSED QUESTION EXAMPLES:**

- Did participants increase their knowledge on the topics covered by the training?
- Did the participants feel confident they could apply the knowledge and skills covered in the training?
- Have judges improved their case management practices?
- Were there unforeseen barriers that prevented the participants from applying their learning?





### **UNIVERSAL Q.3: Are these the right things to do?**

This question is a prompt to consider how effective the training was in contributing to wider impact, the processes that led to this and how it could be improved.

If the training activity did not contribute to the intended impact, then it prompts consideration of why this is the case, and whether unrealistic or unreasonable assumptions have been made about how things would likely happen. It also prompts reflection on whether there are things that happened as part of the training activity that were not expected. This question may also tell us whether any new information or opportunities in the situation have become apparent that could be harnessed to enhance the effectiveness of the training activity in the future.

### **FOCUSED QUESTION EXAMPLES:**

- Did running the training lead to the achievement or an improvement towards the intended impact, such as:
  - An improvement in the volume of cases being processed by the courts?
  - An improvement in access to justice of x social group?
- Is there something else that needs to happen to make progress towards the intended impact?
- Is case management practices the right area to focus on?
- Is it addressing a real need?
- Have better case management practices led to improvements to access to justice? How?
- What has been learnt about the situation (unexpected results, new information, opportunities)?



## ESTABLISHING INDICATORS & TARGETS OF SUCCESS

To provide an evidenced based answer to each of the focused questions, it is important to establish relevant indicators, and targets as thresholds of success. Defining indicators and targets for success will refine how a question is to be evaluated and the kinds of data that will be needed to be collected, which in turn will guide what methods will be best to employ.

Many of these will be quantitative indicators, corresponding to data that can be counted. Quantitative targets for success should be defined in relation to expectations of a project plan, what is achievable and may also need to take into account results specified by donors/ funders or other partners.

Bear in mind that indicators and targets often only

take into account a narrow definition of an aspect of the thing to be measured. For example, to answer the focused question: 'Was the delivery of the training of a high quality?', we may use a 'proxy' indicator of participants satisfaction scores. 'Quality in the delivery of training' may be defined in various ways. Participant satisfaction may be a sufficiently strong indication of the quality of delivery, but is unlikely to capture all aspects that may need to be considered. In some instances it may be useful for questions to be defined by more than one indicator, so that data can be triangulated.

Some focused questions may require an exploration of 'how' and 'why' a target for success was or was not reached. Such indicators will often be qualitative and therefore not composed of data that can be counted and statistically analysed. For instance, collecting stories on people's experiences with the justice system. This qualitative data will require a more interpretive or subjective analysis, since it cannot be compared with a clearly defined target.





## EXAMPLES OF DEFINING TARGETS AND INDICATORS FOR FOCUSED QUESTIONS:

### Q.1: Are we doing what we said we would do?

#### ■ Focused questions:

- ☐ How many judges attended the training?  
(Indicator: number of judges attending, Target: 12)
- ☐ How many female judges attended the training?  
(Indicator: number of female judges attending, Target: 6+ (or 50%))

### Q. 2: Are we creating the difference/change we intended to make?

#### ■ Focused question:

- ☐ Were the learning outcomes of the training achieved?  
(Indicator: Increase in participants' knowledge of 'X', Target: 50% increase (on post-training test compared to pre-training test baseline))

### Q.3 Are we doing the right thing?

#### ■ Focused question:

- ☐ Did the training lead to an improvement in the volume of cases being processed by the courts?  
(Indicator: number of cases being processed by the courts, Target 25% (increase from baseline))



## TIP: ESTABLISHING THE STARTING POINT OR 'BASELINE'

A baseline figure for a particular indicator may be required to be known before the training has happened, to then be able to measure whether there has been a change (such as an increase) in this, due to the training. This may be something that can be researched prior to the training, or established through data collected at the beginning of project implementation.

## PART 2: DATA COLLECTION METHODS

Having focused on understanding what constitutes a successful training activity, the second part of this toolbox will explore different data collection methods to help track and evidence the changes that have occurred as a result of the training. An M&E plan should encompass a variety of tools and methods for the collection of reliable and insightful data.

While it can be challenging to evidence change at the rule of law level and attribute this change to the trainings and programme of one organisation, by utilising a range of data collection methods it is possible to assess both the short- and long-term effects of the training.

The following section of this toolbox is an overview of some commonly used M&E tools which can be employed to collect different kinds of data.

### MEASURING IMMEDIATE RESULTS

***Universal Questions 1&2: Are we doing what we said we would? Are we creating change?***

A good M&E plan would generally use a range of data collection methods to increase reliability and also measure different aspects of the training.

### ATTENDANCE SHEETS

It is important to have attendance records, such as a sign-in sheet, to have evidence of the number and profile of participants. For online trainings, attendance data may be registered and managed through a platform, or participants can be asked to log in with their name in order to capture attendance data.

In conjunction with a system to collect this data, you should keep a well-managed and secure database of training participants with their contact details. This will allow you to keep track of which participants attended which trainings and those who it might be most useful to follow-up with, though you will need to ask for their permission if you intend to do this. It also can evidence how particular participants are staying engaged in a programme of successive trainings and whether their capacity is being built over time, or reflect on what actions or

adaptations need to be made to increase consistent engagement if this is not happening.

## Confidentiality

Of course, there may be situations in which participants are not happy to sign in or provide personal details. For example, if the training is on a politically or socially sensitive topic, or providing such details may be sensitive in itself, such as for undocumented refugees

and migrants. It is worth always assuring participants (and ensuring) that data will be kept securely and not shared with anyone other than the organisation and partner and donor organisations. Participants should always have the opportunity to opt out of collection or storage of their data. In this instance, de-personalised profile data can be recorded about attendees, so long as individuals cannot be identified from this.

## FEEDBACK SURVEYS

ADVANTAGES	DISADVANTAGES
<ul style="list-style-type: none"><li>■ Can collect demographics</li><li>■ Can collect data on the immediate results (e.g. learning outcomes) and views on behaviour change</li><li>■ Can collect quantitative and qualitative data</li><li>■ Produce comparable findings</li><li>■ Quick to administer</li></ul>	<ul style="list-style-type: none"><li>■ Subjective</li><li>■ Can't measure long-term impact of trainings</li></ul>

Feedback surveys are the most commonly used tool for monitoring trainings as it is easy to administer a short survey at the end of a training and/or before the training has begun. But there are many factors to consider that can affect the quality of data you collect through a survey.

## Measuring behaviour change

Feedback surveys can gather information on a range of topics. While feedback surveys are often used to gain quick and simple feedback on participant satisfaction with the training, they can also be used to gain insight into aspects that underpin behaviour changes that could be attributed to the training.

A feedback survey could help to answer the following questions:

- Do the participants feel that they have acquired new skills and knowledge? In which areas (e.g. to carry out a new or existing role)?
- Do participants feel motivated to change their approach?
- What further support do participants need in order to apply their new capabilities?
- Can participants identify ways in which they plan to apply these new skills and knowledge?

## Collecting quantitative and qualitative data

Feedback surveys ought to collect both quantitative and qualitative data. Quantitative data is collected through close-ended questions answered on a scale (e.g. 1-5 or “very poor”, “poor”, “neutral”, “good”, “very good”) or yes/no, while qualitative data is collected through open-ended questions.

Quantitative data is easier for conducting statistical analysis and quick comparisons of results – as everyone is responding using the same measure, and interpretation of the data is less subjective – as well as for creating graphic representations of data, such as graphs. Meanwhile, qualitative data is useful to provide more informative answers and uncover ‘how’ and ‘why’ something happened, if questions are phrased carefully, so as not to be leading.

If the survey is purely qualitative, with people answering open-ended questions, it will be harder to analyse the data afterwards as the participants may not answer in the same way and the evaluator will have to interpret the responses. It will be more difficult to draw generalisations about what happened. It is also more demanding on participants to answer a lot of open-ended questions. On the other hand, if survey questions only yield quantitative data (where people answer on a scale or closed questions), the information will be more limited and not likely to be as insightful to understand how or why things happened.

It is good practice to follow a quantitative question with a qualitative question to better determine the reasons behind the response. For instance, we might know that 60% of participants reported that they gained new ideas from the training, but which ideas were these? This could be as simple as putting a ‘please explain your

response' afterwards. However, such add-on questions often don't elicit thought-out answers or are skipped, so it can be useful to make the question more specific to encourage a response. For example: "If so, what was the most interesting thing you learnt?"

Of course, it is also important to keep the feedback survey concise and proportionate to

the length of the training activity. Most surveys are written to take no longer than five minutes to complete, but could reasonably be longer when evaluating a training that lasted a number of days. It is therefore good practice to request qualitative feedback only where this is of most use, rather than for every question, in order to optimise completion rates.



## TIP: REDUCING SUBJECTIVITY

'The main disadvantage of feedback surveys is that they will only provide subjective data, since answers are based on the participants' own perception of their experiences of the training. One way to reduce subjectivity bias is by asking respondents to provide examples. For instance, "Do you feel able to apply the knowledge/skills gained from the workshop?" could be followed by "Can you give an example of a scenario in which you will be able to apply this new knowledge/skills?". This extension of detail is a corroboration of their initial answer.

It also helps to be as specific as possible in the wording of questions. For instance, instead of "*How satisfied were you with the training?*", you could ask "*How well did the trainers communicate information?*" or "*How engaging was the presentation?*"

This might mean that you have to break one general question down into a few more specific questions but it will give you a lot more useful information and doesn't take much more time to answer if they are asked in the same format, e.g. 4 questions answered on the same scale.'





## FEEDBACK SURVEY CHECKLIST

- Will the questions collect all the information you need? (refer to focused/universal questions)
- Are the question and answer options clear and relevant?
- Are the answer options appropriate to the questions and do they enable participants to express themselves sufficiently?
- If scales are used, are they consistent, e.g. is a 5-point scale used throughout, where 5 is always the best and 1 is the worst?
- Do participants have options to give more general comments/information?
- Are any questions repetitive?
- Does it take more than 5 minutes to complete?
- Is the survey anonymous, while still asking for useful demographic information?





## EXAMPLE TRAINING FEEDBACK SURVEY QUESTIONS

The following are generic questions for training feedback surveys that can be adapted for your purposes.

**1. a. How satisfied were you that the workshop:** *(1 being very dissatisfied and 5 being very satisfied)*

- |                                      |   |   |   |   |   |
|--------------------------------------|---|---|---|---|---|
| ■ Was engaging?                      | 1 | 2 | 3 | 4 | 5 |
| ■ Provided you with new information? | 1 | 2 | 3 | 4 | 5 |
| ■ Clearly communicated information?  | 1 | 2 | 3 | 4 | 5 |
| ■ Went into enough detail?           | 1 | 2 | 3 | 4 | 5 |

**b. Which part of the workshop did you find the most interesting or useful and why?**

**c. Which part of the workshop did you find the least interesting or useful and why?**

**d. How do you believe the workshop could be improved: (please select all that apply)**

- Better facilitation
- More time for participants to discuss concepts
- More time for participants to practise methods
- The workshop cannot be improved
- Other, please specify: \_\_\_\_\_

**2. To what extent do you feel that the training increased your knowledge of the following:**

*(1 being not at all and 5 being to a great extent)*

- |           |   |   |   |   |   |
|-----------|---|---|---|---|---|
| ■ Topic 1 | 1 | 2 | 3 | 4 | 5 |
| ■ Topic 2 | 1 | 2 | 3 | 4 | 5 |
| ■ Topic 3 | 1 | 2 | 3 | 4 | 5 |





**3. a. How likely are you to apply what you have learnt from this training to your work?**

*(1 being very unlikely and 5 being very likely)*

1      2      3      4      5

**b. Can you give an example of how you might apply what you have learnt from this training to your work?**

**c. Is there anything that might prevent you from applying what you have learnt? *(select all that apply)***

- Nothing will prevent me from applying what I have learnt
- Still lack sufficient knowledge on the subject
- Need more time to practise new skills
- Lack of confidence
- I don't think it is an important topic
- The training was not applicable to my line of work
- Other, please specify: \_\_\_\_\_

Note that, in this example, Universal Q.1 is addressed by survey question 1, measuring the perceived quality of activities. Universal Q.2 is addressed by survey questions 1 and 2, measuring knowledge transfer question 3, measuring behaviour change; especially opportunities and motivation to apply knowledge gained in the training. These survey questions are unlikely to provide data directly to respond to Universal Q.3.



## TIP: COLLECTING DEMOGRAPHICS

Feedback surveys and pre/post training tests can be anonymous but still record demographic profile information. Depending on the context and the objectives of the training, relevant demographic information can be gender, region, institution, religion, profession, etc. Such data helps us to better analyse the type of people who are attending and to monitor diversity.

As well as being useful for monitoring equal opportunities in who is being reached through trainings, gathering participant data is very useful once it comes to analysing the data. It allows for pulling apart statistics and understanding who is answering in a certain way, which then may help to understand the barriers being faced by certain groups and determine what could be done to address this.



## PRE- AND POST-TRAINING TESTS

ADVANTAGES	DISADVANTAGES
<ul style="list-style-type: none"><li>■ More objectively measure change</li><li>■ Measure change in knowledge, capabilities and motivation</li><li>■ Measure whether training is the right level for the participants</li><li>■ Identify knowledge gaps</li><li>■ Measure changes in participants' views</li><li>■ Make the training more interactive/ used as knowledge affirming teaching tool</li><li>■ Integrate feedback into the training to increase response rate</li></ul>	<ul style="list-style-type: none"><li>■ Can be laborious</li><li>■ Are still subjective, especially if the questions are self-reflective</li><li>■ Do not explain why knowledge or confidence did/did not increase</li><li>■ Do not explain which aspects of the training worked best or failed</li><li>■ Do not indicate whether participants will use the knowledge they have gained as expected</li><li>■ Some senior professionals may not appreciate being 'tested'</li></ul>

Pre- and post-training tests can provide a more reliable measure of knowledge gain and other changes that have resulted from the training, such as confidence.

There are two ways to design such tests. The first is a self-assessment, asking participants to report on their perceptions of their knowledge and abilities. This is still subjective but enables you to better measure change. The second is more like an actual test (usually multiple choice) including a set of questions before the training that test participants on the knowledge intended to be gained over the course of the training. The same questions are then asked again at the end of the training. Ideally, if the training is set at the right level for the participants, they won't know the answers at the start and will at the end. This therefore provides a more objective measure of knowledge and confidence gain. It also helps you see where there are clear knowledge gaps. For instance, if all participants are still getting a particular question wrong, it suggests that the training is not adequately covering that topic.

Designing these tests is also a useful way to reflect on the training, by establishing the specific knowledge outcomes to be achieved, and whether the design of the training facilitated learning.



## EXAMPLE KNOWLEDGE TEST QUESTIONS

Which of the following is correct?

### **Systemic trial monitoring:**

- a. Is the term used for large-scale trial-monitoring programmes aimed at assessing the functioning of a justice system as a whole in order to support justice reform
- b. Describes programmes with an in-depth focus on one or several thematic areas, such as war crimes, administrative justice or pre-trial proceedings
- c. Focuses on the observation of individual, usually high profile, cases or groups of cases
- d. Not sure



## EXAMPLE SELF-ASSESSMENT CONFIDENCE QUESTION

**On a scale of 1-5 (1 being not at all confident and 5 being extremely confident), how confident do you feel that you:**

- Understand the six primary objectives of trial observations?

1      2      3      4      5

- Are able to put the six primary objectives into practice?

1      2      3      4      5

- Can write trial observations to a high standard?

1      2      3      4      5

Designing these tests is also a useful way to reflect on the training, by establishing the specific knowledge outcomes to be achieved, and whether the design of the training facilitated learning.



## TESTING DATA COLLECTION METHODS

It is useful to ensure there is time for colleagues to review feedback surveys and pre- and post-training tests before they are used in the training, to help ensure that questions are clear and easy to answer.

For pre- and post-training tests, it is important to check that the answers cannot be easily guessed and are set at the right level.

## OBSERVATIONS

ADVANTAGES	DISADVANTAGES
<ul style="list-style-type: none"><li>■ Intuitive</li><li>■ Undemanding</li><li>■ Not dependent on the action or responses of participants</li></ul>	<ul style="list-style-type: none"><li>■ Less objective</li><li>■ Harder to compare findings</li><li>■ Do not tell you why things happened</li></ul>

Observations are a useful tool since they don't ask anything of participants and we intuitively do this.

Everyone naturally observes things that go on around us and may make judgements and speculate why things are happening. But to conduct observations for M&E purposes, natural observations need to be shifted up a gear to become more analytical.

An observation template is a useful tool to

have to help make focused and systematic observations about relevant things. It should prompt reflections on certain aspects of the training that otherwise might be overlooked or not well considered. The template can be adapted to be relevant to each training and M&E plan.

It is good practice, when taking notes, to include anecdotal evidence explaining why certain conclusions were made. For instance, to support an observation that participants were not



engaged, it is worth also reflecting on reasons why no-one answered the trainer's questions.

However, though it is possible to improve the reliability of any observations made, it is still hard to identify the reasons behind these

observations. While observations are useful for providing a general impression, they remain subjective, and therefore cannot be relied on as the sole method for gaining robust data. Rather, they should be used in conjunction with other data collection methods to triangulate findings.



## EQUALITY MONITORING

Having a seat at the table doesn't mean having a voice. Equality monitoring is as much about observing equal interactions, such as who is asking and answering questions, as it is about recording attendance.





## TRAINING OBSERVATION TEMPLATE

Observer: \_\_\_\_\_ Date: \_\_\_\_\_ Training: \_\_\_\_\_

1. Which organisations/actors were present at the meeting?
2. Were there any notable absences in terms of relevant stakeholders or expertise?
3. Was the strategic importance of the training clear?
4. What was the added value of each expert/speaker?
5. How did the training incorporate issues of gender equality and social inclusion?
6. What aspects of the training session were most effective, and why?
7. Did you notice any way that the training could be improved?



QUESTION	YES / NO / N/A	NOTES (DETAILS, EVIDENCE)
<b>DID THE TRAINER...</b>		
State the purpose of the training and learning outcomes clearly?		
Explain how the training is relevant to participants' job/role?		
Check participants' current knowledge/ experience?		
Follow the training plan?		
Explain information clearly and in detail?		
Create a participatory learning environment (e.g. suitable power dynamics, positive learning opportunities...)		
Give participants enough opportunities to ask questions?		
Encourage all participants to engage in learning?		
<b>DID THE TRAINING...</b>		
Cover all the important points?		
Include activities to reinforce new knowledge?		
Provide the opportunity to practise new skills?		



QUESTION	YES / NO / N/A	NOTES (DETAILS, EVIDENCE)
<b>WERE PARTICIPANTS...</b>		
N.B. Make note of any differences observed between the behaviour of different groups of participants, such as by gender or profession		
Engaged in discussions?		
Asking questions that suggested the training was set at the right level?		
Providing answers to questions that demonstrated learning?		
Enthusiastic about what they were learning?		
<b>LOGISTICS</b>		
Was the venue appropriate? If online – was the internet connection strong enough? Was the software used suitable?		
Was the timing appropriate? (Did the training run on time? Was enough time provided for participants to arrive? Were there enough breaks?)		
Were all the necessary training resources available?		
Was the training accessible to participants with a disability?		

## MEASURING IMPACT

### *Universal Question 3: Are these the right things to do?*

When we talk about impact, we mean the long-term effects of the training (and surrounding activities) in reaching the intended goal - for example, an improvement in the rule of law, and subsequent improvements to people's lives.

It can be difficult to measure impact when working on rule of law and other development issues, as the desired change usually takes several years to be realised. Furthermore, change will be the result of several factors, rather than a single training activity, or the work of one organisation alone. Therefore, in most cases it will not be possible to determine a simple 'X' training directly led to 'Y' scenario.

However, there are still ways to monitor if the training is contributing to longer-term change, such as if case outcomes have been improved or relevant legal reforms have been achieved.

## PART OF A BROADER M&E STRATEGY

Methods for measuring impact are not only for measuring the impact of trainings but for measuring the impact of the combination of activities that form a project/programme and

how the training contributed to achieving the intended end goals. They should help to assess if the programme strategy is working - testing the assumptions in the plan or theory of change and identifying any unexpected effects of the training and other project activities.

### **FOLLOWING UP WITH PARTICIPANTS AND OTHER STAKEHOLDERS**

The simplest way to know if the training led to a long-term behaviour change that has led to wider impacts is by following what the participants go on to do. This may be through a follow-up survey or interview sent several months after the training – leaving enough time for changes to have occurred but not long enough for participants to have forgotten what the training was about - and consider how the training and other activities have affected their behaviours.

The questions that are asked may be about a range of activities that participants have been involved in, not only focused on a specific training. In addition, speaking to participants a number of months after the training will help to identify how different elements of an activity or project plan have improved their skills and behaviours over time.

At this point, it is very helpful to have kept good records or a database of participants, as well as having built strong relationships with the training

participants and other stakeholders involved. The advantage of an interview (or catch-up call) over a survey is being able to ask follow-up questions. However, if there is a large number

of participants, it's advisable to send a survey round first, and then follow-up with those that yield interesting results if it would be valuable to get more information.





## **EXAMPLES OF QUESTIONS TO STRUCTURE SURVEYS AND INTERVIEWS AROUND INCLUDE:**

- Has the participant remembered the information delivered in the training? Do they need additional resources?
- Since the training, has the participant had a chance to use the knowledge gained? If so, what were the details of the outcomes? (the more information here, the better)
- Have they encountered any obstacles or challenges in doing so? What could help to address this?
- Do they now feel that there is anything important that was missed out of the training?
- How have the different activities (if asking about a range of activities) supported their capacity-building? Which were the most useful?

## OUTCOME HARVESTING

Sometimes it can be hard to measure impact by following developments from X to Y because we can't realistically keep following up on all participants to see how the training is affecting their behaviour and ways of working, and whether this has then resulted in institutional and societal change. Instead, we might need to identify the

change (Y) and then work backwards to see if the training and other activities can be evidenced as having played a role in creating this change.

This is a method called outcome harvesting that focuses on collecting evidence of change and then working backwards to determine whether and how specific activities have contributed to that change.



### OUTCOME HARVESTING IS MOST USEFUL UNDER THREE CONDITIONS:<sup>iv</sup>

1. When measuring change in the long-term (i.e. impact on behavioural, social and institutional change), as opposed to only in the short-term (such as number of trainings carried out, the number of people engaged and learning outcomes achieved).

2. In complex situations where the

relationship between cause and effect is not fully understood and/or where many different actors influence change.

3. When stakeholders want not only to identify change, but also to learn about how and why those changes were brought about to improve future performance.

While you may require external M&E expertise at the analysis stage of an outcome harvesting approach, organisations ought to be responsible for collecting data to assist this analysis. Essentially, you are collecting the pieces of a jigsaw that an external evaluator can then put together to tell the story.

Types of data collection that can assist in an outcome harvesting approach include any evidence of changes occurring at a higher level. These depend

on the project's end goal, but may include evidence of such changes in the conduct of fair trials in line with international standards, changes to legislation and increased public trust in the legal system.

Evidence of such changes may be found in:

- Legal reports
- Media reports/press releases
- Interviews with key stakeholders
- Research reports and surveys



## EVIDENCE BOXING

A good system for storing such data is 'evidence boxing'. This is as simple as having a folder (or folders) online or on a shared network where all the evidence of any of the type of changes mentioned above is stored. All team members should be briefed on what to look out for and practices in how and where to store it.







## DATA COLLECTION METHODS OVERVIEW

This provides an overview of the kinds of data that can be collected using different data collection methods and how the data corresponds to answering focused questions that would fall under the three universal questions for M&E.

METHOD	UNIVERSAL Q. 1 ARE YOU DOING WHAT YOU SAID YOU WOULD?	UNIVERSAL Q. 2 ARE YOU MAKING A DIFFERENCE/ CHANGE?	UNIVERSAL Q.3 ARE YOU DOING THE RIGHT ACTIVITY?	TIME AND EFFORT
Attendance sheets	Who is attending	N/A	Helps with follow-up	Quick and simple
Feedback forms (participants)	Questions can be designed to respond to this	Subjective measure of elements of learning and behaviour changes	Doesn't give us information on the long-term effects/impact of the training	Relatively quick
Pre- and post-training tests (participants)	N/A	Objectively measures change in capability and motivation	As above	Longer to design, quick to complete
Observations (trainers/ members of facilitating organisations)	Observations can confirm whether the activity meets Q.1, but only 'internally' and subjectively	Provides indications of observable changes, but requires validation	As above	No effort for participants, more effort for trainers/ organisation members
Follow-up with participants (surveys and interviews)	N/A	Can help validate actual behaviour changes	Can determine the impact of behaviour changes	Requires a commitment to make contact and chase participants
Interviews with key stakeholders	N/A	Can help validate actual behaviour changes	Can help to determine impact of behaviour changes	Takes time to contact stakeholders
Outcome harvesting	N/A	Can help validate behaviour changes	Can help to determine impact	Data collection isn't time-consuming if processes are in place



## ADAPTING YOUR STRATEGY FOR VIRTUAL ONLINE TRAININGS

While running your trainings virtually shouldn't affect your M&E plan too much, there are some important things to consider.

### ATTENDANCE SHEETS:

It is easier to forget to collect details of attendance for online trainings because you can't place a sign-in sheet at the door. It's important to remember to make a record of final participants, which will require requesting participants to sign-in with an identifiable name. Alternatively, platforms such as Zoom, include report functions that record logins, but will not collect other profile or personal data.

### FEEDBACK SURVEYS:

One significant challenge that online trainings pose is the completion of feedback surveys, since it is harder to encourage participants complete these before leaving 'the room' and completion rates in asking participants to complete surveys after is often low without an additional incentive.

Ways to improve completion of feedback surveys include:

- Design the training so that the feedback surveys are completed as part of the training and before closing discussions.

- Explain why the surveys are important for both the trainers and the participants.
- Integrate surveys into the conference call software so they automatically open up in another window when the training ends (though this means that they have to be completed after the training).
- Use a live virtual poll as part of the session, where results will be visible to participants and can prompt further discussions.
- Send reminder emails a couple of days after the training if there is a low feedback rate.

### DIGITAL RECORDINGS:

There can also be advantages to running trainings online. For instance, it may be possible to record sessions for sharing with others who may not have been able to attend. This record may be used to review the session and make further observations.

### INTEGRATED TOOLS:

Some platforms also include a single or multiple-choice polling function. However, such functions are typically not as adaptable for M&E purposes as specialised survey software/websites are.

Using online survey software can also save time that would have been used for data entry, as results can be downloaded in a variety of formats, including as a spreadsheet (such as in Excel).



It can also facilitate the data analysis process, as reports are automatically generated. Be careful though that the data has been 'cleaned' – this means checked for errors. Automatically generated surveys don't take account of errors such as if one person responded twice.

Lastly, if new to running trainings online, remember to include an assessment of the online format of the training in your M&E plan. For instance - Were participants still able to interact well? What had to be left out of the training and how might this affect the results?



## M&E STRATEGY - PLANNING TABLE

Considering the list of focused questions and indicators/targets, will establish exactly the data you need to capture. From this, and with consideration of the opportunities that the training activity will provide, effective decisions can be made about the most appropriate methods of data collection for your strategy.

By completing the M&E strategy planning table, you will determine for each method:

- What approach is going to be used and which questions it will aim to answer
- What the sources of data can be (the people or documents/resources)

- The preparations required before it can be delivered (and by whom)
- How it will be delivered (and by whom)
- How the data will be recorded/stored (and by whom)

The strategy should also confirm what the data will be used for and how this will be managed, e.g. to upload the data into a spreadsheet and for this to be relied upon to draft a report or to provide feedback during a future planning of the project.

You can ensure that the strategy is relevant by checking that all your focused questions are incorporated into the methods and referring to the principles of M&E.



36

# PART 3: EVALUATE AND ADAPT

## DATA ANALYSIS

The data collected is only useful if it is analysed. The time and resources needed for this should be built into M&E planning.

It is common that stages of data collection and analysis will need to be repeated. This is because the analysis of data may give insights, that then illuminate gaps in understanding about what has happened, or the need to understand 'why' things have happened. This may prompt follow-up surveys or interviews with participants to investigate further any unexpected results or issues that were missed in the first round.

### DATA ANALYSIS STEPS

#### Step 1: Consolidate and organise the data

After data collection and before data analysis, the data will need to be consolidated. This means that raw data, for instance from surveys, needs to be organised into usable formats, such as within tables on spreadsheets. Online survey tools typically provide spreadsheets ready to download.

#### Preparing data for processing:

'Cleaning' data – This means removing any clear errors in the data by correcting the entries, if you are certain of the answers, or removing the data if you are not certain.

'Coding' data – This is the process of organising and labelling data to give it a clearer meaning. For example, that a respondent has answered '1' on the form – which refers to the response 'excellent'. Another example is that it may make sense to combine the numeric scores of several responses from a participant, to create an 'index' score – which may provide a more useful statistic for the data.

#### Step 2: Create useful 'descriptive' statistics

Many of the targets that have been established against your original focused questions for your M&E plan will require the calculation of simple statistics. These are known as 'descriptive statistics' from the data which describe its general characteristics, such as counts of results (e.g. 25 women attended the training) or percentages of total respondents (90% responded that they feel 'extremely confident' to use what they learned).

Additionally, using mean averages of scores of data, such as the average response to a question about confidence (from a sliding scale 1-5) can give general insights about the training and what changes occurred, comparing before and after data.

It is sometimes useful to calculate percentages for quantitative data. When citing percentages of something, it is important to note how many the total is, for example, to note the number of participants that percentage statistic represents. For example, 60% (12 participants) reported that they found the training 'very useful'.

A more complex statistic being a calculation of knowledge gain – this requires calculating the increase of scores, between before (the baseline) and after the training. This might be presented as the real change (in terms of total average scores), or as a percentage change, between the baseline score, and the after score.

For example, if the average baseline (before) score of 14 participants attending the training captured through a knowledge test is 3.5/10, and the average after score is 8.5/10 – the percentage change can be calculated as follows:

$8.5 \text{ (after score)} - 3.5 \text{ (baseline score)} = 5 \text{ (the change/increase in average scores)}$

$5 \text{ (the change)} \div 3.5 \text{ (baseline score)} = 1.42,$

which represents an increase of average scores by 142% of the average baseline score.

It is also important to report the number of respondents that the data has been collected from, and what percentage this number is of the total number of participants who attended the training. The larger percentage of the total number of participants that have provided data, increases confidence that the data as representative of all participants.

### **Step 3: Summarise qualitative data findings**

Rather than transcribing all the qualitative data collected, it is acceptable to provide a summary of what has been said, while indicating how many people answered a certain way. The summary should help to reveal patterns in the responses, rather than distort them. Sometimes it may be possible, and useful, to suggest an analysis of what the findings are showing, but suggestions should be clearly highlighted as such in the report.

The presentation of qualitative findings should be integrated into the report in a way that provides further insight into the statistics. The training report template below suggests a way to structure the presentation of findings.



#### **Step 4: Disaggregate data – to understand experiences of different groups**

As part of your analysis, it is important to disaggregate data to identify any differences between the experiences of, and effects of the training on, different groups of participants. The ways in which you are able to disaggregate data will depend on the demographic information you gathered at the time of data collection. Typical variables for which to disaggregate data include gender, profession, and location.

It can also help provide more insight into findings to disaggregate one set of answers based on the answers to a different question. For example, you might want to identify the number of people who said that the training had improved their knowledge and also said that they would be able to put this knowledge into practice. This is simple to calculate if there is a spreadsheet of the data but is, of course, more

time-consuming if you are making calculations directly from individual survey forms.

#### **Step 5: Compare data to increase confidence in findings**

To increase the reliability of your findings, it is important to triangulate data – that is, cross-checking data collected through different methods or sources. Triangulation helps to verify findings, as well as identify inconsistencies. It is particularly important to triangulate qualitative data. For instance, when conducting interviews, consulting a number of stakeholders allows you to compare different people's accounts.

## **REPORTING FINDINGS**

It is useful to consolidate your findings into a report.

You may want to structure your report by answering the three Universal Questions of Monitoring and Evaluation and the focused questions that you generated when planning your M&E plan.







## TRAINING REPORT TEMPLATE

Begin the report with general information about the training event and the sources of data for the analysis, such as number of survey respondents and interviews conducted, including relevant demographic information. Include any important reflections that support the strengths or limitations of the data being relied upon.

### Q1: ARE WE DOING WHAT WE SAID WE WOULD DO?

- No. of participants and participant demographics
- What was delivered – training contents and methods (including a note of any planned activities that didn't take place and any unplanned activities that took place)
- Who was involved in preparing and delivering the training?
- Quality of the training outputs – questions to consider here include:
  - Were the methods used effective?
  - Was the content at the right level and included all necessary information?
  - Were the trainers appropriate?
  - What was the added value of having pro bono experts involved?
- Sources of information:
  - Quantitative findings from feedback surveys, such as % of participants satisfied with the training
  - Qualitative data from feedback surveys and observations, such as feedback on which parts of the training they found the most and least useful, suggestions for improvement, etc.

### Q2: ARE WE CREATING THE DIFFERENCE/CHANGE WE INTENDED TO MAKE?

- Have participants acquired new skills and knowledge from the training?
  - Relevant quantitative and qualitative data from pre- and post-training tests (including: % knowledge gain), feedback surveys, observations and follow-up calls

- Are participants motivated to change their approach?
  - Relevant quantitative and qualitative data from pre- and post-training tests (including % confidence gain), feedback surveys, observations and follow-up calls
- Are participants likely to apply these new skills and knowledge?
  - Relevant quantitative and qualitative data from feedback surveys and follow-up calls
- Do participants need any further support in order to apply their new capabilities?
  - Relevant quantitative and qualitative data from feedback surveys, observations & follow-up calls

## ADAPTING TRAININGS FROM THE FINDINGS

Data analysis is primarily a way to support evidence-based learning about a training activity, the wider project or programme. Therefore reporting on M&E findings for individual trainings, as well as at important milestones in the project, provides the basis for making evidence-based decisions about the project going forward.

A report of findings should be shared with other relevant stakeholders and used to inform debriefing sessions. Relevant stakeholders are those directly involved in the delivery of the activity, but could include others such as training participants.

Debriefing sessions help to determine and agree on any adaptations to the training in the future. It is a way to consult on different perspectives and utilise the findings in the M&E report to robustly evidence the basis of any decisions made to adapt the activity.

The structure of the debrief can reflect on answering the three universal questions of M&E and the focused questions. The debrief should also reflect on the effectiveness of the M&E processes and any adaptations that may be needed to improve them. If it was found, for instance, that the data collected was not able to answer all of the universal or focused questions, or that there is uncertainty about the findings, then the M&E approach and tools may need adapting.

# CONCLUSION

This toolbox has been conceived as a practical guide to comprehensively designing, planning and implementing effective monitoring and evaluation methods that are tailored specifically for training activities.

We hope that the toolbox demonstrates well the value of applying considered approaches and techniques for monitoring and evaluation. These methods should help to know how to best adapt and improve trainings and more widely development intervention projects aiming at strengthening the rule of law.

Integrating monitoring and evaluation methods at the very beginning of the design and planning of trainings will ensure that M&E strategies are effective without being onerous. Undertaking M&E will improve the quality of your trainings and ensure that activities are making a real difference to strengthening the rule of law.

If you have any comments or specific needs in terms of further monitoring and evaluation guidance for your training, please contact us at **[roleuk@a4id.org](mailto:roleuk@a4id.org)**.



# FURTHER RESOURCES

## ■ Project Cycle Planning and M&E

### ROLE UK: Monitoring & Evaluation Toolbox

Legal Assistance for Economic Reform. [The why, what and how of monitoring and evaluation: guidance for providers of international pro bono legal assistance](#)

Project Management for Development Organisations ([pm4dev.com](http://pm4dev.com)) has a comprehensive list of manuals and guidelines for use in development project management on its site: <https://www.pm4dev.com/resources/manuals-and-guidelines.html> \_

[Toolkit for Monitoring and Evaluation Data Collection](#), Pacific Women Shaping Pacific Development. This includes guidance on data collection methods and analysis, including

survey design.

For useful templates and a how-to for developing project indicators, visit [Tools4Dev](#)

## ■ Outcome Harvesting

[Outcome Harvesting](#), INTRAC, 2017

[Outcome Harvesting](#), Better Evaluation

<https://outcomeharvesting.net/home/>

## ■ Data Analysis

More details on analysing qualitative and quantitative M&E data can be found at [Measure Evaluation](#)

Step 3 of the International Federation of the Red Cross and Red Crescent Societies' [M&E Guide](#) also offers useful tips for data analysis

# NOTES

<sup>i</sup> See <https://www.roleuk.org.uk/resources/maximising-technical-assistance-improve-rule-law-guide>

<sup>ii</sup> See <https://www.roleuk.org.uk/resources/monitoring-and-evaluation-toolbox>

<sup>iii</sup> See <https://www.roleuk.org.uk/resources/monitoring-and-evaluation-toolbox>

<sup>iv</sup> Taken from Outcome Harvesting, INTRAC (2017), <https://www.intrac.org/resources/outcome-harvesting/>

