Firdale Consulting

HOW TO CREATE AN M&E SYSTEM

Firdale Consulting

About us | We are a team of economists and development professionals that specialises in data management and quantitative monitoring and evaluation for impact investors (foundations and corporates), non-profits and research organisations. We work with clients to establish change pathways, write code to routinely clean data (Python, R and Stata), leverage existing data for analysis, and provide the technology needed to monitor and evaluate impact.

Founded in 2016, we have worked with clients across a wide range of sectors including early childhood development, education, HIV/AIDS, public health and the green economy.

Our vision is to work with complex data for good, equip non-profits and funders with tools for data analysis, and contribute to making monitoring and evaluation more rigorous.

TABLE OF **CONTENTS**

INTRODUCTION

This guide outlines the requirements for a comprehensive monitoring and evaluation (M&E) system for non-profits, donors and government agencies. Developing such a system can be undertaken in a seven-step process: 1) understanding the needs and constraints, 2) developing the conceptual tools, 3) collecting and cleaning data, 4) designing the system, 5) automating reports and developing a dashboard, 6) identifying roles and assigning responsibilities, and 7) conducting M&E training. While this guide is comprehensive it is not intended as an M&E handbook and expert advice should be sought when implementing M&E solutions.

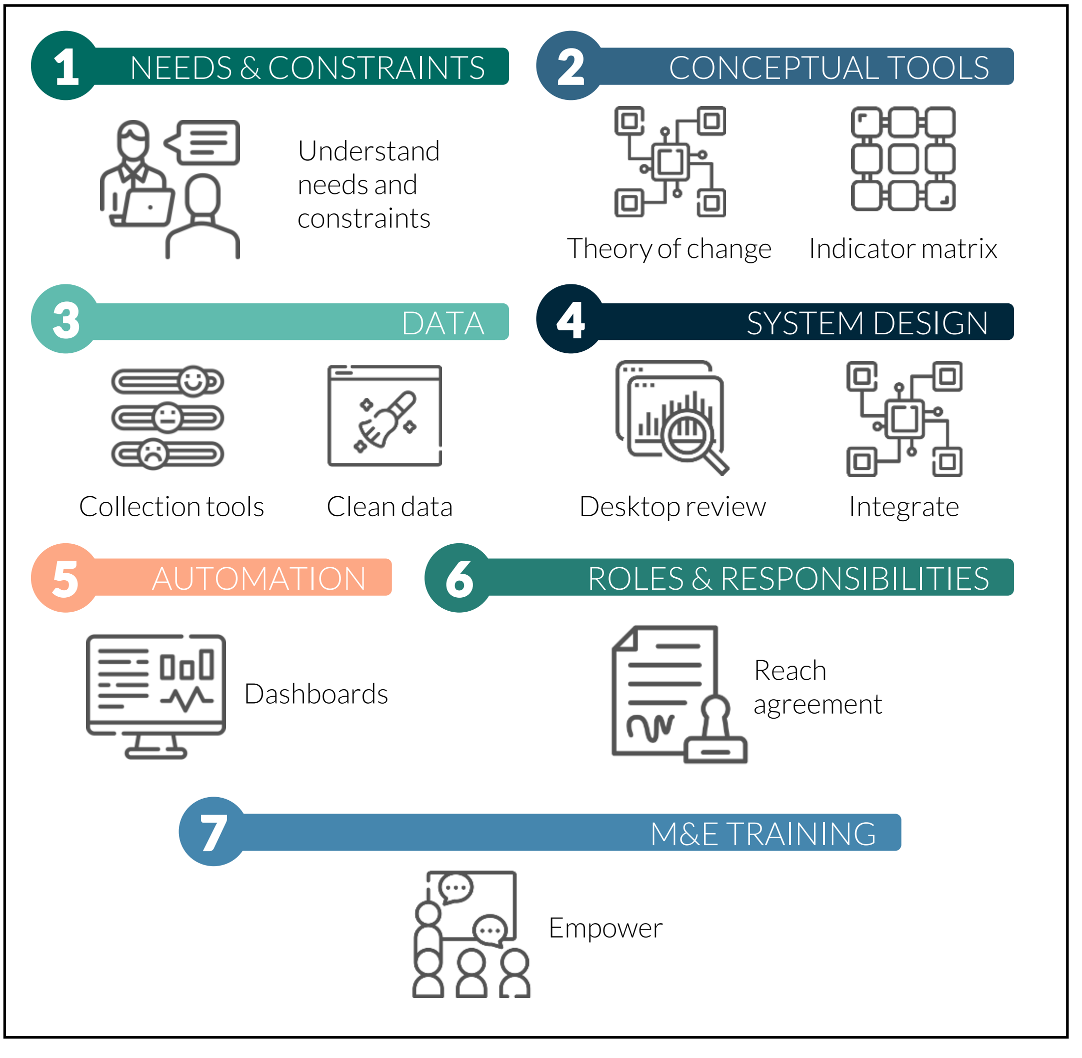

Figure 1: The seven building blocks of a comprehensive M&E system

1. NEEDS AND **CONSTRAINTS**

To ensure the M&E system is practical, useful and sustainable it is important to understand what needs the M&E system is intended to meet and the constraints that will be faced.

Understanding needs

It is critical to understand what each organisation or 'implementer' expects from the M&E system. Often organisations want to track progress towards their core mandate, or they want insight into "blind-spots" where they are struggling to measure progress or make decisions. Thus, at the start, clarify whether the programme's performance can be measured using a data-centric M&E system, and if so, what type of system is best. It is recommended that a discussion of general needs (typically facilitated by an M&E expert) is had around the following M&E talking points:

Context / mandate area

What key areas does the implementer need to monitor, and is it feasible that an M&E system can help? If so, do we know of any analogous or comparable existing systems in this area that we can model our discussion around?

Presentation

What kind of presentation (tables, figures, visualizations or even in written word) does the implementer have in mind, and is it appropriate? Visualizations and tables can be interactive and more advanced systems can allow the user a level of 'customization' over the features.

Output complexity

How complex should the output be? It is possible to simplify information into key statistics or numbers, and present these using simple visual tools or infographics. It is also possible to create more technical visualizations or tables, showing statistical distributions, confidence intervals, marginal distributions etc.

Timing

How often does the implementer need to consult the M&E system, present results or update the underlying data? This is also contingent on the data review (discussed below) and the frequency at which the available or new data can be collected.

1. NEEDS AND **CONSTRAINTS** (Cont.)

Reporting format

A major output of M&E systems is report generation. It is critical to understand what kind of reporting is expected: to whom are reports given, how often and what are the key points? What format should the document and the information be in? It is also likely that reporting and other structures are already in place, especially if the organisation is part of a broader group.

Targets

It is usually necessary to know and display the targets and/or baselines for the measures we are interested in. What are the timelines and units of the targets and can they be measured with existing data?

Data collection

What kind of data are expected to be collected? Does the project require the collection of new primary data and if so, how will this be done? Are we sure the data do not exist somewhere else already? If not, is there a comparable survey or data collection project we can model our data collection project on?

Data cleaning

Is it necessary for the M&E system to be fully automatic or does the implementer have capacity to clean new data as they come in? Data cleaning is discussed in detail below.

Flexibility

To what extent does the implementer expect to be able to interact with and change the M&E system? For example, do they want to be able to add new data sources, or change the reporting format?

1. NEEDS AND **CONSTRAINTS** (Cont.)

Understanding constraints

There are a variety of constraints that face an M&E system which typically include finance, skills and technology.

Finance is an important constraint: how much funding is there to develop this system and how is the funding budgeted across the various stages of implementation? This affects the scope of system design, as well as the capacity to collect and clean data. For instance, projects that want to collect their own data are naturally more expensive, but this can be mitigated with cost-reduction strategies, such as using cheap or open-source data collection software or leveraging existing resources (for example enumerator's smartphones. Typical costs include designing, building and maintaining the data system, training staff to understand and use the system, and the time and skills of the staff involved.

Next, it is important to consider the skillset of the implementing team. M&E systems that are 'owned' by the implementing team often work the best and so taking into account the team's skills regarding technology (for instance programming new surveys), theoretical knowledge (designing effective survey questions), and managerial skills (managing fieldworkers for instance) is very important. A system that is too complex for the context is less likely be successfully adopted.

Technological constraints are also relevant to the M&E system. For instance, if WiFi is unavailable solutions must allow data to be collected in an offline context for later upload. As another example, the implementer might be part of an organisational network and the system needs to draw and collate data from systems of other organisations. This would require obtaining permissions, understanding the technology they have employed and potentially leveraging APIs to access their data automatically.

Figure 2: Reviewing the constraints

The goal of M&E is to provide insight into the impact of a project and whether its objectives are being achieved, in a quantifiable and objective way. However, before one can measure success, it is necessary to understand what 'success' means, and this requires mapping the goals of the implementer. Two fundamental components help us achieve this – the Theory of Change (TOC) and the Indicator Matrix (IM).

Theory of change

The TOC is a logical diagram that helps an organisation visualise how it will pursue a desired impact. The TOC is comprised of various components which are laid out to show how they relate and combine to support the end goal.

TOCs typically contain 5 components: inputs, activities, outputs, outcomes and impact. These components are sequential – we start with inputs and mix these to create activities, activities create outputs and outcomes, and these in turn lead to the intended impact. The figure below is a generic representation of this process:

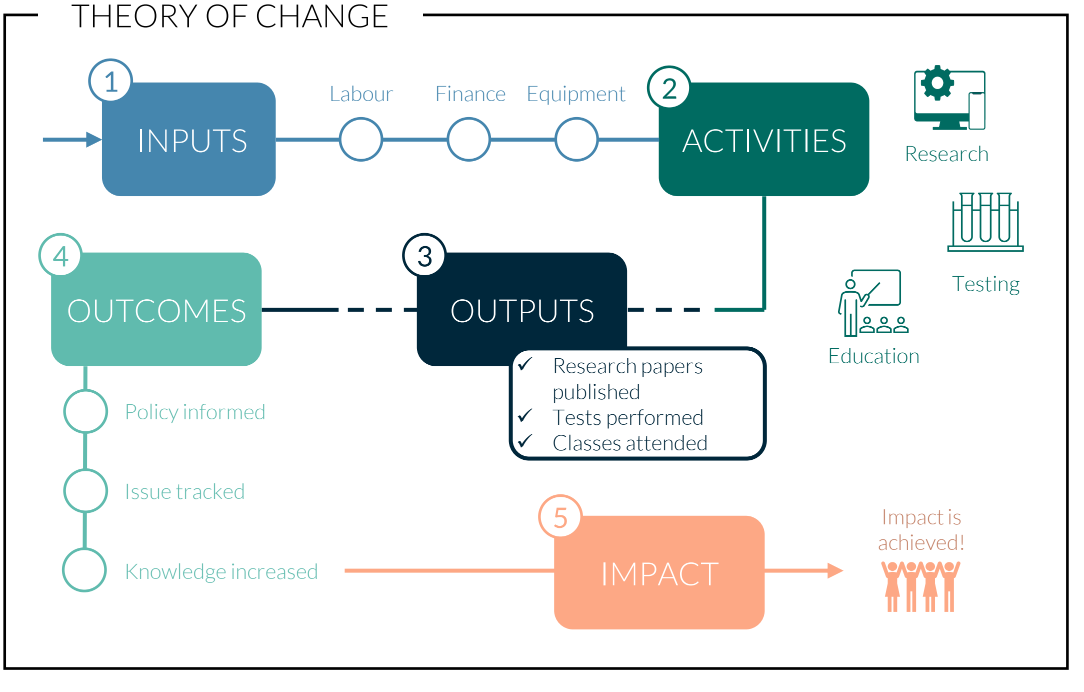

Figure 3: A theory of change links an organization's inputs to its intended impact

2. CONCEPTUAL **TOOLS** (Cont.)

Explaining these 5 components is best done by example. Let's say we are mandated to reduce the spread of HIV– this means our intended impact would be the reduction of HIV. Our **inputs** are the resources we can combine to try combat HIV; typical examples are money, time and equipment. These we combine to produce **activities** (or 'interventions' as they are also known), which take place in the real world. Possible activities might be to test for HIV or to run HIV awareness campaigns.

If the activities are successful they create outputs and outcomes. **Outputs** are quantifiable units the activity directly creates – for example the number of HIV test results produced or the total views of the adverts run by the awareness campaign. **Outcomes** on the other hand are the benefits that these outputs are intended to create (for instance an increased proportion of people knowing their HIV status would be an outcome of the testing activity). Lastly, the **impact**[1] is the end goal that is being pursued and this is something the organisation would contribute to but not expect to achieve on their own.

Laying these components out in the TOC helps to make clear the causal links we expect between them and how these contribute to our outcomes and ultimately the desired impact. Note that there is not always a one-to-one relationship between the components. For example, several activities can have the same output, or a single activity could create many outputs. The same is true of the relationship between outputs and outcomes – there are many possible combinations.

Because the TOC helps to reveal everything we should expect to see, it is indispensable to comprehensive M&E. It informs us as to what we should measure for and can also help reveal where the system is breaking down, for example if an expected outcome is not being achieved as expected. Importantly, the TOC is a live document that should be updated as required over time.

1 Another way to understand the distinction between outputs and outcomes is that outputs are the means, whereas outcomes are the ends. Outputs are "neutral" items or events such as an HIV test or a condom being used, whereas outcomes are the ends of these, being knowledge of HIV status and increased protection during sex.

Indicators

Once we understand the components of the TOC we need to measure whether these are being achieved. Indicators are the tool we use to do so.

Indicators are measurement tools that can be quantitative or qualitative. A quantitative indicator measures something that can be enumerated (counted), and it is usually given as a number or a percentage. Quantitative indicators tend to be technical and precisely defined, for example "the number of HIV tests performed and returned in a week". Qualitative indicators on the other hand are used to examine things that are not strictly enumerable or numeric. They are therefore broader and allow us to capture more complex, nuanced or loosely defined concepts.

Again, an example will make this clear. Let's say we want to know how easy people find it to access HIV clinics. There are several possible quantitative measures we could devise, for example: "how far is it to the nearest clinic from home?", or "how long did you wait to be attended to at the clinic?". However, even if we use many such measures, we might miss something important or focus on the wrong areas. This is because "easiness" is a complex concept that differs depending on context. In this case a qualitative indicator might be useful, for example something like "how easy have you found it to access HIV clinics in the past year?".

Qualitative indicators can therefore be useful as they help us capture broader concepts. However, it is important to make sure indicators do not become meaninglessly broad. Developing precise indicators – whether they are qualitative or quantitate – makes them more valuable. The SMART framework is a useful way to make sure we are developing high-quality indicators.

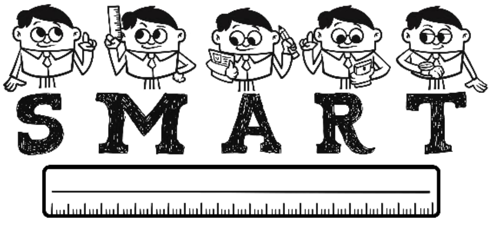

Figure 4: SMART indicators

2. CONCEPTUAL **TOOLS** (Cont.)

A SMART indicator is specific, measurable, achievable, relevant and time-bound. Let's take a look using the "distance to nearest clinic from home" indicator to see how it can be improved using SMART:

S

Specific means the indicator leaves little-to-no room for interpretation or ambiguity. We could improve the distance indicator by specifying that it is the distance *from your primary address* to the nearest clinic *offering HIV-related services*. Naturally, the less room there is for interpretation, the more consistent the indicator will be across observations, time and space.

M

Measurable means the indicator has a well-defined system of measurement including the units and technique of measurement. For example, our distance indicator should specify that it is measured in kilometres, perhaps as a straight line on the map from home address to clinic.

A

Achievable means the indicator is practically viable. Imagine if we had instead specified measurement as "the distance along the shortest path by car"; while it might be more accurate, the technology to calculate it would be incredibly expensive.

R

Relevant means the indicator is logically related to all or part of the input, activity, output or outcome we are trying to evaluate. For example, in a context where most visits are in person and travel is costly, distance is highly relevant to the overall ease of accessing HIV services. However, if some of the services are offered online (such as HIV counselling), distance becomes less relevant.

T

Time-bound indicators make clear the relevant time-frames. Do we want to know how easy access was over the person's lifetime, in the past year, or the past month? This not only makes the indicator more specific but also informs us on how often we might want to collect data. If the question asks about the past year, it makes sense to collect it annually.[2]

[2] For a blog post with more information on SMART indicators please see: https://www.firdaleconsulting.com/blog-post-2

Photo by Kaleidico on Unsplash

Specifying indicators

An indicator that meets the SMART criteria will have a few constituent parts and it can be useful to separate these out rather than have them in one long run-on sentence. UNAIDS[3] recommends that a single indicator be defined with the following separate parts: a title, a definition, a purpose, a rationale, a method of measurement, and where appropriate a numerator and denominator. When we are dealing with many indicators it can also help to give each indicator a unique code.

Let's take the example of an indicator that measures the average distance to the nearest clinic:

Indicator Code	b1-dist.
Title	Average distance to nearest clinic.
Definition	The average distance from individuals' primary residence to the nearest clinic offering HIV-related services.
Purpose	This indicator is one of a handful of indicators developed to measure how easily people can get access to HIV-related services.
Rationale	The HIV services provided at clinics are critical in reducing the spread and harm of HIV. However, one of the major deterrents or obstacles people face when getting HIV services is the costly commute to a clinic. Therefore, understanding the average distances people face is crucial in evaluating the ease of access as well as the prevalence of HIV services in general.
Method of measurement	The straight line distance is measured using GPS coordinates of households and their nearest clinics.
Numerator/ Denominator	We take the total of all the distances and divide by the number cases.

Figure 5: Example of a well-specified indicator

3 See the guide *An Introduction to Indicators* (2010, UNAIDS).

Indicator matrix

It is helpful to organize indicators in an indicator matrix (IM). In essence the IM is a table that lists the indicators and all of their components. Each row is a new indicator and the columns capture their constituent parts (the bold items from the figure above).

The IM helps make sure that the activities and outputs in the TOC have associated indicators, and that each indicator is properly specified. Generally, activities and outputs are focal in the IM and this is because they are quantifiable, making them more practical to measure.

The IM can also be used to categorize indicators and there are a few ways to do this. One system is according to the items on the TOC – indicators can be grouped as activity, output or outcome-related. Indicators closer to our end goal (i.e., our impact) are often called impact indicators. The term coverage is sometimes used to group together indicators that measure the extent to which intended effects are reaching across the target population. A custom category can also be used for indicators that are unique and do not fit the other categories. Ultimately the grouping depends on the logic of the overall project.[4]

Another important consideration is whether the IM should cover all conceivable indicators or if it should be limited to only those that are practically attainable given the constraints of the project. Lastly, there are often standard sets of indicators that are used in a given field; incorporating these into the IM where relevant is recommended because the results will then be more comparable with other work in the field.

---

[4] For example, in the HIV context a useful grouping is the three categories of combination prevention services: biomedical, behavioural and structural.

3. **DATA**

Once we have completed the conceptual tools we have a good idea of what we need to measure with the M&E system. Next, we move on to working with the data. This happens in three steps: review, collection and cleaning.

Data review

The data review is a preparatory step before data collection where we perform a 'census' of all data sources that could potentially be used for the indicators. The review should make clear the access conditions of each possible data source (public, private, licensed, etc), how well the variables in the data file(s) match the indicator, the frequency at which the data are updated, and the amount of processing that will be necessary to get the data into the correct format. Note that some references suggest doing the data review with the IM and not as a separate step. This is a choice of the M&E developer.

It is likely that during the review it is discovered that some indicators cannot be easily measured as the required data are not available. In this scenario there are three options: 1) look for a proxy measure of the indicator, 2) collect new data or 3) leave the indicator as "desirable but currently unobtainable". If the IM was done separately and prior to the data review, it is generally better not to go back to the IM and change or remove the indicator because it is still a relevant part of the overall strategy, even if it cannot be measured at present. When using proxies it is important to work with analytic experts that can conduct the necessary analysis to generate accurate proxy indicators.

Data collection

Data collection is the process of gathering data. Data collection can be done in different ways, including: 1) compiling data from publicly available existing data sources, 2) requesting data access to existing data sources that are not publicly available, 3) conducting desk research and compiling data from secondary sources such as written reports, or 4) creating and collecting primary data yourself by designing and executing interviews, surveys and other data collection instruments. The collected data should be stored in a single repository such as a physical or virtual server.

Creating quality data (if applicable)

Importantly, if the project requires option (4) from above - the collection of primary data - all necessary steps should be taken to ensure the data are of sufficient quality. A data collection project that creates valuable data must begin by following international best practices in survey design, questionnaire design and software deployment.

Good survey design will ensure that the final data provide a meaningful sample that can be used to represent the intended population, rather than a set of observations and variables that just so happened to be collected. By contrast, happenstance and poorly thought-out sampling leads to analytic problems due to sample selection biases or incorrect/invalid sampling weights, for example.

A well-designed questionnaire will be easier and quicker to complete and is capable of significantly reducing the number of errors captured. When used correctly, collection software can be designed to identify, reduce and remove errors before data collection begins. This is done by pre-coding answers/responses to the questionnaire using a **codebook** and then specifying the correct or valid answers for each question. Wrong answers (such as an age of 150 years) can be automatically stopped before data are captured – a process called **input validation.**[5]

Figure 6: Data cleaning code

5 For more on survey, questionnaire design and related topics, see our blog: https://www.firdaleconsulting.com/blog

3. DATA (Cont.)

Identifying and resolving errors early in this fashion is important because errors become exponentially more difficult to repair the later through the process they are detected, a concept we return to later. Another way to check for problems early in the process is to run a pilot survey before the main fieldwork begins. During a pilot survey we get a small group of test cases to answer the questionnaire and then check what the data look like. This is an important stress-test of questionnaires which might reveal errors or problems with the questionnaire, as well as potential problems with the enumeration team.

Checking is another crucial strategy for error mitigation when you are collecting your own data. There are three types of check: high-frequency, spot and back. High frequency checks occur often (for example daily or weekly) and can begin as soon as data collection does. Analysts check the data to see that fieldwork is on schedule and that enough observations are coming in, that the distribution of the respondents makes sense, or that the distribution of their responses seems reasonable. Because these checks can start on day one of fieldwork, they have the important benefit of being able to spot mistakes early on. A spot check is when a higher-level team member (with expert knowledge) checks in on enumerators at random to see that they are conducting the interview correctly. This might also involve looking at a handful of the enumerator's previous completed questionnaires.

Lastly, back-checking is when a respondent is re-interviewed a second time by a new enumerator and asked some of the same questions as before. If large discrepancies arise that is worrying and needs to be investigated. A downside to back-checking is that it can only occur after we have several completed surveys, by which time it may be too late to correct the problem if it affects some/all the other interviews.

Creating quality data (if applicable)[6]

As touched on above, errors become increasingly expensive to fix the later they are detected. Therefore when collecting data it is wise to spend relatively more resources on developing a sound data collection system, rather than trying to fix problems after the fact. A rule of thumb that illustrates nicely this is the 1-10-100 principle displayed on the next page.

[6] For an excellent resource with more detail on data checking see JPAL's article here: https://www.povertyactionlab.org/resource/data-quality-checks

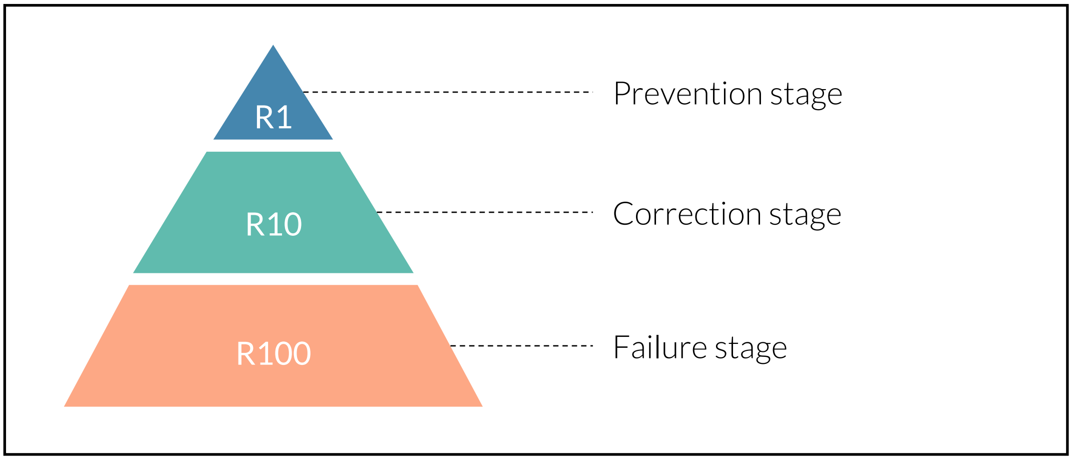

Figure 7: The 1-10-100 principle illustrates the increasing cost of errors

The prevention stage is during system design, where we use the tools discussed above to plan for and treat the possible errors that may arise. If the error is not prevented, it is passed to the correction stage, when the data have been collected and passed to a cleaning/analysis team. If the data cleaning team them cannot resolve the error, it passed to the failure stage. The failure stage is when the error remains in the data and ends up having a negative effect on the quality of analysis.

As the figure makes clear, the cost rises exponentially the longer the error remains. By example, programming software to check that the values are correct is very cheap. However, if we fail to do so, the hours spent in data cleaning going through all the wrong values and correcting them is roughly 10x as expensive. Worse, if we fail to clean the incorrect answers, and the data then get used for analysis, the potential cost (in terms of wrong results and wrong decision making) is incredibly high, possibly 100x that of input validation that could have been done in the prevention stage. This is why good system design is critical.

Data cleaning

The 1-10-100 principle illustrates the general point that “prevention is better cure”, in other words avoiding errors is better than trying fix them. However, no matter how carefully we design systems, errors do arise. This is because it is impossible to predict everything that might happen: respondents in unexpected responses, enumerators make mistakes, and all sorts of practical problems occur from weather events to changes in the political climate.

3. **DATA** (Cont.)

Data cleaning is therefore an important part of an M&E system. Although data cleaning is primarily used to address errors and the cost thereof, it also involves certain other quality-improving techniques. In general, any strategy that alters data in order to improve its use-value can be considered data cleaning. To ensure that data cleaning is as free **from human error** as possible, **scaleable** as the data collection grows and **replicable** by anyone, it is important that data cleaning be conducted through computer code, using statistical programs such as Stata, R or Python. Whilst some organisations may be tempted to conduct data cleaning through a more manual process (for example cutting and pasting or typing corrections into Excel), this inevitably introduces errors and increases costs over time, especially as the project grows.

Below we discuss some of the intricacies of data cleaning and recommend employing experts in data cleaning to streamline the processes. The natural process by which to clean errors is: 1) look for an error, 2) attempt a fix if feasible, 3) check if the error is resolved, and then start again at 1). This process is summarized below:

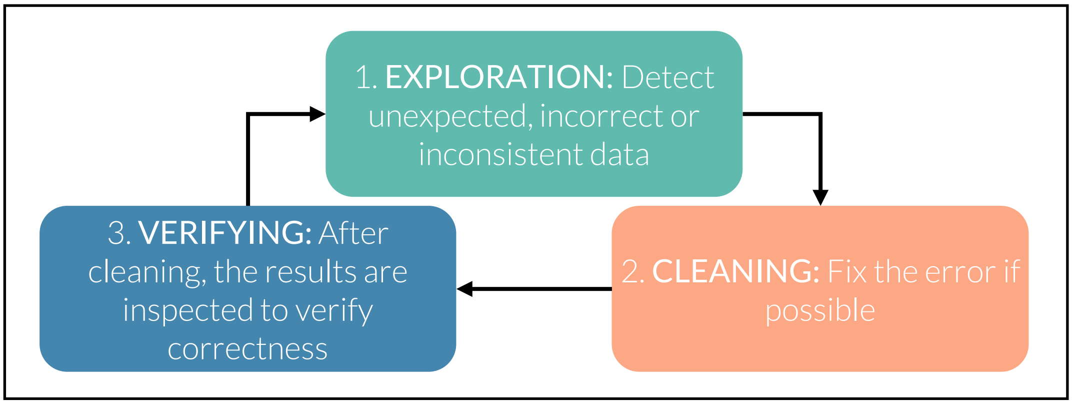

Figure 8: Iteratively attending to errors to improve data validity

There are many errors we might encounter. We might find values out of range, data of the wrong type (e.g. text instead of number), suspiciously large or small outliers, self-contradictory data, observations out of sample, observations that look like duplicates, observations with problematic (non-distinct or unexpected) ID, etc. While attending to errors a minimalistic attitude is advised. In general, it is better to avoid editing the data where possible, for two reasons: 1) it empowers anyone analysing the data after cleaning to decide how to handle the errors and 2) doing so creates potential for the data cleaner to introduce their own errors. This is especially true if the cleaner is making assumptions about what the data were "meant to be", as these assumptions could be wrong.

More complex improvements can also be made if the data cleaners have the relevant skills: weights can be adjusted, languages translated, geographic data reprojected, categorizations can be improved or converted via correspondence, and so on. In all cases, if changing the data is deemed necessary it is important to follow **transparency** principles. All changes to the data should be meticulously documented, as doing so improves transparency and assists in metadata creation.

In addition to improving validity, data cleaning can be done to improve **efficiency**. We can convert categorical variables to numeric storage using encoding, which is more computationally and storage efficient. String and numeric variables can be shortened (without any data loss); removing redundant or completely erroneous/empty variables also improves efficiency. This becomes increasingly important as data continue to be collected over time and the datasets continually expand.

We can also adjust the data to make it more **user-friendly.** Common calculations such as creating totals or splitting a variable into a grouped version (such as aged groups) can be done in the data cleaning phase, so that each subsequent data user does not need to do it again. Datafiles can be properly named so that they are easy to identify and the relationships between them clear. If variable names have not been pulled successfully from the questionnaire, variables can be renamed following a prefix-root-suffix system which helps users navigate the data. If the dataset involves data files that can be combined (for example, hierarchically, as a panel, or as a time series), test runs of the combining and the results thereof should be published with observation counts, so that future users can replicate and confirm that they get the same results.

Photo by Sajad Nori on Unsplash

4. SYSTEM **DESIGN**

1 2 3 4 SYSTEM DESIGN 5 6 7

Once we have the conceptual framework and the data, we can build the M&E system, which will take in the data and produce the required reports and output, either automatically or semi-automatically. System design requires two steps: a desktop review of the possible solutions and then the integration of the new system with existing systems.

Desktop review

A desktop review is conducted to research and evaluate software solutions that can be implemented to create various kinds of M&E systems. In the desktop review we examine relevant documentation about M&E system development in the given context and look for solutions.

The following decisions should be focal. Firstly, is the project trying to save costs using open-source software, or is proprietary software an option? What key functionality does the M&E system require and which packages can achieve this? Does the software necessitate permanent connectivity? What software do the organisation's partners use? In a nutshell, the desktop review aims to find software that meets all the requirements we have documented. There are a plethora of M&E software options out there and it is useful to review each with experienced practitioners that understand their pros and cons.

Integrate with existing systems

The next facet of system design is how the new software will integrate with existing tools. It is crucial to understand what systems are being used by:

- The M&E implementer
- The organisations collecting input data
- The organisations using the output or intermediate data

Leveraging and integrating with these will ensure that the system runs smoothly and has good uptake. This includes the existing ways in which data are collected, stored and exchanged, the existing reporting channels and tools, and the software used to this end. Implementors that exist in a broader organisational structure or bureaucracy will already have data sharing and reporting systems in place. They may already be part of larger M&E systems, data/information management systems and application programming interfaces (APIs). Tapping into these as data sources and software tools will reduce duplication of effort and training costs.

5. AUTOMATE **REPORTING** AND **DASHBOARDS**

A good M&E system will generate a visual interface in line with the user needs. This could include automated reports, automatic WhatsApp messages or a dashboard (amongst other options). Provided the data system is fed updated data in the correct format, the above functions can be run whenever necessary to create new and updated output, by simply running the program. Any dashboarding elements that have been designed (indicators, summaries, tables, visualizations) should also update.

The first step is to prepare a "mock-up" or template version of what dashboard and reporting will look like. This can be done iteratively as the project develops, so that adjustments can be made in light of any new needs, constraints, software considerations or other relevant information. Next, the design is implemented onto the chosen software. After that, the system is trouble-shooted (tested for errors and usability) by multiple users and all issues are addressed. It is useful for the M&E system designers to be available for any system updates over time – for instance if an API changes or new functionality is needed.

Figure 9: Automated report' Source: Firdale Consulting

6. AGREE ON M&E **ROLES AND RESPONSIBILITIES**

Once the M&E system is up and running the developers will be able to demonstrate how it works. At this point it is crucial that everyone understands their roles and responsibilities.

M&E systems typically require maintenance. For instance, input files may need to be manipulated into the correct format. Similarly, if there is a change to one of the data sources (such as the addition of new variables or a reshaping of the file) figuring out how to incorporate the new changes can be non-trivial. The system itself can require updating, for example if the software it has been built on goes through a major update. Compatibility can be an issue if the system is running over various software packages.

It might later transpire that the system needs improvement for certain reasons. One of the core functions might not be performing due to an oversight in development, such as an omitted variable that is needed for disaggregation. It is possible that the deployment of the system has revealed that some quite major additional tools are needed.

It is therefore important to identify the various roles that will be required and assign responsibility for these – either within the implementing organisation or with an external expert.

7. M&E **TRAINING**

Finally, all relevant users of the M&E system should be provided with training and a manual on how to use the system. This should include various aspects of the system – from data review, collection, cleaning, visualisation and reporting.

Figure 10: Zoom workshop training

In addition, for many organisations maintaining an M&E system can feel daunting and M&E training can help to empower organisations with the necessary skills. The scope of the training will depend on the existing skills as well as the needs of the organisation. Training can cover any to all of the concepts that make up an M&E system: the conceptual tools (TOC and IM) and the concept of SMART indicators, survey design and execution, data cleaning, data analysis, system use and maintenance, etc. Perhaps most importantly, training should include demonstrations of how the system works.

Training can happen using a host of tools from in-person workshops to virtual lectures on Zoom. Firdale Consulting can also run short-course "bootcamps" on data cleaning software and analysis techniques if necessary. Key individuals identified in section 6 should be targeted with the knowledge most relevant to them. In general, the more thorough the training is, the more likely it is that the system will be a success in the long-run.

CONCLUSION

1
2
3
4
5
6
7

CONCLUSION

We can close this discussion by directing the reader to the Anna Karenina principle. The principle was originally developed as an anecdote to describe successful families. It can however be translated into the M&E context to provide insight. Translated to this context it would read: "All successful M&E system are alike; each unsuccessful system is so in its own way". What the principle means is that success can only occur if the key set of ingredients are mixed in precisely the correct way; if they are not, all sorts of wrong results can happen.

M&E systems follow the Anna Karenina principle. Each of the seven steps is equally important in ensuring that the system is functional, and all sorts of strange mistakes, bugs, workarounds and outright failures will occur if components are neglected. It is therefore important to cover each thoroughly.

ABOUT THE AUTHOR

Bruce McDougall (M.A. Economics) is a data analyst with a healthy obsession for data quality. He is expert in data management and analysis across multiple sectors including education and health. Bruce has helped to prepare and document several of the most popular research datasets in South Africa, and has published and presented papers on these topics at expert conferences.

Contact Bruce at:
bruce@firdaleconsulting.com