

Outcomes Measurement TOOLKIT



TRUE CHARITY®

Why Measure Outcomes?

Poverty-fighting nonprofits exist to see individuals move from poverty to flourishing. We will never know if our programs accomplish this unless we measure results.

Instead of only measuring how much we give away, we should measure long-term results like stable housing, employment, education, and family reunifications.

With the right tools and program design, it is possible to measure even intangible things like knowledge, motivation, and social connectedness. If we're not measuring, we run the risk that our programs are ineffective or even counterproductive.

Quantitative measurements don't replace personal interaction and anecdotes, they augment them. If you're buying a product online, you may read a few customer reviews to find details about the product and how people liked it. You will also check the average customer star rating of the product to ensure the reviews you read are representative of what most people thought. The numbers and the stories together tell you far more than either could alone.

There are many types of things you can measure in nonprofit work. All the measurements in the categories below are valuable. However, this toolkit focuses on the last type of measurement, outcomes measurement. This is the most challenging, most overlooked, and most critical to your mission.

Any use of this material constitutes an agreement by the user to abide by all Watered Gardens and True Charity copyright terms. True Charity®, Watered Gardens®, Project Worth®, Life Deck®, and Neighbor Connect® are trademarks of Watered Gardens, Inc. All rights reserved. Our full copyright disclosure can be found on the True Charity website at truecharity.us/copyright.

Table 1: Types of Measurements

Type of Measurement	Definition	Examples	Purpose
Financials	Measurements related to funding and stewarding your resources	Donations, expenses, donor retention rate	To sustainably fund and grow your program
Demographics	Details about the clients and their needs and capabilities	Client age, employment status, family size, education levels	To understand your clients and tailor programs to their situations
Inputs	Resources required to produce outputs	Number of volunteer shifts filled, number of classrooms needed	To manage the resources required for your program
Outputs (Activity)	A program's activity and its countable products	Number of clients completing courses, meals served, nights of shelter	To keep track of your organizational efficiency in delivering services
Outcomes (Impact)	Long-term changes for those served by a program, generally in knowledge, skills, attitudes, behavior, or condition.	Number of clients who obtained full-time work, number of clients who retained the work after six months	To ensure your programs work and you are accomplishing your mission

"Most of the giveaway volunteer efforts initiated by religious and charitable groups are working at cross-purposes with programs committed to moving people out of poverty. Such giveaway programs consider their existence proof of their success; they hardly ever evaluate their success in terms of moving the poverty needle."

ROBERT LUPTON, AUTHOR OF TOXIC CHARITY

Here are three main reasons to measure our outcomes:

1) IT ENSURES OUR PROGRAMS WORK.

In most cases, charity work takes place across a divide of cultures, values, backgrounds, and even languages. In this context it is easy for a program that seems sensible to a staff member to not work in practice. Furthermore, if clients leave and your program has no positive impact on them (or just perpetuates dependency), they are not likely to come back and inform you. You may have no way of knowing unless you make a conscious effort to collect the results.

2) IT GIVES SUPPORTERS A REASON TO SUPPORT

Donors and volunteers (particularly major donors) want to know that their generosity is making an impact. As they have entrusted us with their gifts, we owe it to them to prove their dollars make a difference. In return, they will be likely to increase their involvement. Granting organizations especially want to see your organization's outcomes. Even paid staff will be encouraged by demonstrable impact. It is likely that they are making a sacrifice to work in your organization as well.

3) IT SUPPORTS CLIENT PERSONAL GROWTH

Rather than being a merely administrative task, good outcomes measurement can directly support clients' personal growth. Personal assessments used in data collection lead to natural goal setting conversations and clients will be able to see their own positive change over time. For many of them, this is the first time in years (if ever) that the trajectory of their life has been positive, and seeing the improving numbers on things like credit score or weeks of sobriety gives them the confidence to believe the progress is real. Additionally, in life as in sports, people play differently when they are keeping score. Tracking progress and goals in a concrete way is an important way to inspire people to do their best.

How to Measure Outcomes

We will begin with a clarification of terms and a high-level overview of the process we will explore.

The table below is a “logic model.” A logic model charts the logical connections among your mission, resources, programs, and objectives. It is an excellent way to ensure you have clarity about your strategy as well as a good way to explain it to other stakeholders (see table 1 for definitions of inputs, outputs, and outcomes). Subsequent sections will go into more detail on each of the components.

Table 2: Sample Logic Model

Mission Why we exist	Inputs What's required	Outputs What we provide	Outcomes What matters long term
We exist to help people find fulfilling employment.	<ul style="list-style-type: none"> ▪ Classroom ▪ Curriculum ▪ Teachers/ Coaches ▪ Partner employers 	<ul style="list-style-type: none"> ▪ Work Readiness Training ▪ Internships ▪ Job hunt coaching ▪ Job fairs 	<ul style="list-style-type: none"> ▪ Stable employment ▪ Job satisfaction
Example Measures (Indicators)			
N/A	<ul style="list-style-type: none"> ▪ Volunteer coaching slots filled 	<ul style="list-style-type: none"> ▪ Number of people attending class ▪ Graduation rates ▪ Number of employers at job fairs 	<ul style="list-style-type: none"> ▪ Demonstrated increase in job search knowledge ▪ Percent of clients getting jobs ▪ Percent of clients maintaining jobs for six months ▪ Average client-reported increase in job satisfaction

Note on terms: There is no completely standardized set of terms to describe this process and there are a lot of minor variations on the logic model. For simplicity, we will attempt to use the fewest number of new terms possible.

Here are the steps to measure outcomes for your organization.

1. Identify Your Outcomes (and group them into Outcome Domains)

The first step is to identify your desired outcomes—the positive things you hope for in the long run. Depending on the complexity of your program, you should identify 5-15 outcomes. This number of objectives could be difficult to manage, but generally, similar outcomes can be grouped together by category. We call these buckets of similar outcomes “outcome domains.” Typically, 1-4 outcome domains are sufficient for a program; more than 5 is likely to stretch your focus too thin to make much progress.

Example

You run a community outreach program and identify the following long-run goals (outcomes) for your clients:

- Stable employment
- Improved incomes
- Increased net worth
- Elimination of chemical addictions
- Healthy lifestyle
- Close friends and family
- Volunteerism

You could group all these goals into two outcomes domains:

- 1) Financial Stability: stable employment; improved income; increased net worth
- 2) Wellness: elimination of chemical addictions, healthy lifestyles, close friends and family, volunteerism

Or you could create an alternate set of outcome domains such as:

- 1) Personal Freedom: increased income, increased net worth, elimination of chemical addictions, healthy lifestyles
- 2) Connectedness: stable employment, close friends and family, volunteerism

In our experience, there are a limited number of sensible outcome domains in poverty resolution, although terminology varies from place to place. [Click here](#) for a list of a wide variety of sample outcomes, the domains that contain them, and possible ways to measure them (discussed in step 3 in more detail). This list contains the most common outcomes but is not intended to be exhaustive. Additionally, we have categorized the outcomes into sensible domains, but it is possible to recategorize them in other ways as shown above.

Team exercise: If you want to bring multiple people into the process of selecting outcomes, have each contributor write his or her ideas on sticky notes (one outcome per note). Then use a whiteboard or table to group like outcomes (and duplicates) together into outcome domains.

2. Ensure Your Outcomes Connect to your Outputs and Mission Statement

Once you identify your outcomes and group them into domains, you should ensure they logically connect to your organization's mission statement and outputs. The outcomes are your destination, the outputs are the road to get there. *You should always understand where you want to go before you choose a way to get there.*

If you are a new organization, this step will help you determine a mission statement and sensible outputs. **If you have existing programs,** but have never identified clear outcomes before now, you should go through this process to ensure your existing mission and outputs align with your newly defined destination.

CHOOSING YOUR OUTPUTS

Start with your desired outcomes, taking each domain in turn, and then ask "*What's required for that to happen?*" multiple times, until you arrive at clear actions your organization can take that make sense given your clients and resources. These actions might be full programs or parts of programs which we will refer to as outputs. By asking this series of questions for each outcome domain, you can identify which programs are likely to work. This works well as a verbal exercise, but for demonstration purposes, we will provide a written example below.

While choosing programs that are sensible for your needs involves factors such as client demographics and organizational capacity, the exercise of identifying suitable outcomes is an excellent first step to reveal what programs make sense.

Example

Our desired outcome domain is "Social Connectedness."

- 1) What is required for social connectedness to happen? Clients must be *willing* to seek and maintain friends and acquaintances.
 - a. What is required for them to be willing to seek and maintain...? Clients must *understand* the value of friends and acquaintances.
 - i. What is required for them to understand the value...? Clients must interact with people who genuinely care about them and demonstrate the value. [**Possible output:** volunteers or staff deliberately connect with and engage clients in meaningful relationships] OR
 - ii. What is required for them to understand the value...? Clients must be instructed in the value of social connections. [**Possible output:** a one-on-one conversation with case managers who explain the value of social connections in escaping poverty]
- 2) What is required for social connectedness to happen? Clients must have opportunities to make helpful connections. [**Possible output:** create an opportunity for people to make connections outside of their socio-economic class]
- 3) What is required for social connectedness to happen? Clients must be reunited with existing friends and family members (whose influence is positive).
 - a. What is required for them to be united with existing affiliations? Clients must be challenged to connect with them. [**Possible output:** Case managers should incorporate questions about positive family and friend connections and facilitate reconnection when possible as a part of the goal setting process.] AND
 - b. What is required for them to be united with existing affiliations? Clients must have their contact information. [**Possible input:** Case managers should have access to a software subscription that assists in looking up people's contact information given their names, known previous locations, etc.]

"Feeling good about what we do is not a valid measure of the good we do."

JAMES WHITFORD, CEO OF TRUE CHARITY

You can integrate the elements that result from this series of questions (the inputs and outputs bolded in the example above) into a number of sensible programs to promote social connectedness. Depending on your demographics and resources, possible programs that could incorporate these elements include case management, formal mentorship programs, informal cross-group activities designed to build connections around fun or common interests, a club fair/networking event, a rotating in-home dinner hosting schedule, or many more options.

For existing organizations, many times this process uncovers that current programs are not well suited to achieve desired outcomes.

Examples

"Financial Independence" may be a newly identified desired outcome, but the current case management program may be signing all clients up for various government subsidies which they could feasibly get by without. Such a practice would be moving them in the opposite direction of the stated goal. While someone's long-term goals may benefit from temporary financial assistance, a generic practice of signing everyone up for long-term benefits would move many people away from financial independence. This program would need to be modified to avoid signing people up for long-term government benefits when there are other feasible options.

A charitable car garage could have a goal of empowering low-income people to maintain their vehicles in good condition. If their primary program is free or discounted car repairs, this alone is inadequate to equip people to maintain their own vehicles. Furthermore, it could work against that outcome if clients feel like their access to free car repair relieves them of the burden of maintaining their own property. While discounted car repair could remain a central part of this organization's programs, to achieve their goal, they will also require education programs as well as a rigorous application process that ensures the program doesn't merely subsidize poor maintenance.

In general, many organizations are built around transactional relief models (such as distributing a free good or service with minimal interaction), when relational development models are better suited to their desired outcomes. To reach most desirable outcomes, you will need to consider developmental models. If you want to explore some alternate program models, check out our [Model Action Plans](#), [Recommended Client Classes and Programs](#), and [Models of Charity Articles](#).

CHOOSING YOUR MISSION STATEMENT

A good mission is memorable, explains what you do, and guides action. While there is a natural tradeoff between clarity and brevity, a mission statement no one remembers is unlikely to guide action. So, err on the side of memorable.

Sometimes you will realize that your current mission statement does not connect logically to your desired outcomes.

Example

Current Mission Statement: "We exist to eliminate hunger and hardship in our community through the distribution of food."

Desired Outcomes Domains: Career development, Financial Stability, Social Connectedness

Conflict: It often happens that charities open with the idea that people in poverty only need access to free things. Many eventually learn by experience that root problems run deeper than they first thought and that clients are often in greater need of development, accountability, or access to opportunity than access to free things. At this point, the original mission statement no longer connects to the outcomes they want to support.

More Suitable Mission Statement: "We address root causes of hunger and hardship in our community by empowering people through relationships and opportunities."

3. Determine how to measure your outcomes (indicators)

Identifying desired long-term change is inadequate on its own. You must also determine how to measure it. We use the term "indicators" to refer to the outcome measurements themselves. The indicators are the evidence that we have achieved the desired outcome. There are many possible indicators of the same outcome. Take for example an outcome of debt reduction. Possible ways to measure this include:

- Self-reported amount of total debt
- Self-reported debt as a percent of income
- A verified credit score
- Number of payday/title/unsecured loans outstanding
- The pre and post intervention score on a financial literacy quiz about debt related knowledge

You do *not* need to measure all the possible indicators for an outcome. Pick the indicators which are most reliable, simple to collect, and easy to analyze. You will often find that a single indicator can help gauge progress towards multiple outcomes across one or more domains. For instance, the indicator "percent of clients using a written budget" could support an outcome of debt reduction and an outcome of self-management.

[This link](#) provides a list of common outcome domains, the outcomes they contain, and possible indicators. It is not intended to limit your creativity, but it is an excellent starting point.

Consider the following best practices as you craft your indicators.

BEST PRACTICES IN OUTCOME MEASUREMENT

Measure your outcome directly, if possible, indirectly if necessary.

Sometimes it is possible to measure an outcome directly such as measuring a person's total amount of debt or the number of unsecured loans they hold. Sometimes, this may be infeasible and so you can measure something closely related, such as credit score. Occasionally, this may not be possible as well, so you can measure something more indirect, such as a change in knowledge about debt and financial instruments. Direct measurements are preferred, but often you can get creative and find indirect ways to measure.

Measure objective indicators (facts) if possible, subjective indicators (feelings) if necessary.

An objective indicator is something which is a matter of fact, a subjective indicator is a matter of general feeling. More objective indicators typically provide more reliable information. Feelings can vary widely from moment to moment, causing random variation in survey results. "How many social events did you participate in last month?" is a more objective question than "do you feel connected to people in your life?" Sometimes, subjective questions are the only tool available, in which case, try to make the question as specific and factual as possible. For instance, "How many people would you feel comfortable calling in a personal crisis?" is a better question than "do you feel connected to people in your life?," because it is more objective (even though it is still a matter of personal opinion).

Measure primarily quantitative indicators (numbers) and some qualitative indicators (descriptions).

Most outcome indicators can be measured numerically. Even feelings can be rated on numbered scales (1 to 5 is common). Quantitative data is best for reporting, since you can

easily average and graph it. However, there is also an important role for qualitative data such as questions like "What would you tell someone else who was considering joining this program?" or "What part of this class would you improve and how?" These types of qualitative questions can both provide valuable feedback and quotations to share with donors and other stakeholders.

Here are some common formats for quantitative indicators.

Table 3: Formats for Quantitative Indicators

Format	Explanation	Example	Where to Use
Objective Measurement	Verifiable numbers, not matters of opinion	Number of square feet of living space per occupant.	Whenever possible; these measurements allow comparison across clients
Subjective Scale	This is a scale of numbers, often translated from a survey with words (e.g. "very," "somewhat," or "not at all")	On a scale of 1-5, how adequate is your housing for your needs?	When objective measurements are not feasible and you want to compare a client to himself over time
Binary Answers (could be subjective or objective)	These are choices with only two options (e.g. Yes/No translated into 1 or 0—which will allow you to average the results)	Does client have adequate housing? (1/0) Does client own home? (1/0)	Use when scales are infeasible, or would overcomplicate the issue

Measure verifiable indicators if possible, self-reported indicators if necessary.

In formal research, self-reported data is widely regarded to be the least reliable sort of data. On some occasions, you may be able to use data that is not self-reported (such as attendance logs, employer reported information, or health assessments). This is nearly always preferable to self-reported data. However, when you use self-reported data, at least do your best to decrease the likelihood that people misrepresent information (often accidentally). Ensure that instructions are clear, people do not stand to gain from

misrepresenting their situation, and that people feel either relatively anonymous in their reporting, or they feel like they are reporting to someone who cares about them and with whom they have a relationship.

Measure indicators for 100% of your clients when possible, a random sample when necessary.

Sometimes program structure or size doesn't allow measurement for all the clients—or you are trying to get information on clients after completion of a program and not all respond. In this case, you may be able to measure a sample of clients and extrapolate the results. For instance, 100 clients complete job training and you follow up with 25 six months later. If half have stable employment, you can reasonably infer that half of all of your clients (50 of the 100) have stable employment at six months.

However, this only works if the sample you measure is random and representative of all of your clients. For instance, if you attempt to contact all of your clients and only 25 respond, it is likely that the people who respond are different from the people who refuse to return your calls. They may be more likely to be working and proud of their situation.

While you may not be able to overcome this bias completely, you can mitigate it. It is better to reach out to only 25 people, but then go to great lengths to get a response from most of them (calling the last known employer, visiting home addresses, etc.) than to reach out to 50 and take the first 25 responses that come in. Also, ensure the demographics of the people you sample are representative of the rest of your clients. Factors like age, income source, gender, and education level might cause one segment of your clients to respond differently to a program than others—so make sure the sample has similar characteristics to the whole clientele.

COMMON QUESTIONS AT THIS STAGE

What if your program design doesn't support outcome measurement?

A common objection to outcome measurement is that the program in question can't support much measurement. For instance, 100 clients come in one night a week over the space of two hours to receive some service or gift. Typically, programs that don't support outcome measurement are also programs which involve relatively little client contribution and minimal relational connection. For the same reasons that make outcomes difficult to measure (lack of relational connection and client buy-in), these programs also tend to be quite ineffective at achieving much long-term change. Even if you could solve the measurement problem by magic, you would likely be disappointed with the results once you saw the measurements.

However, if program change is not an immediate option, you may be able to integrate measurement by focusing on a smaller sample of clients as explained above. By measuring some fraction of clients every time, you could eventually assess most of them.

How much information is appropriate to ask clients to share?

There is a tradeoff between the intensity of your services and the amount of information clients can reasonably be expected to share. More intensive programs like residential rehab, transitional housing, or school scholarship programs can ask for a significant amount of personal information. Programs that are less intense, like a cooking class or a [Christmas market](#) cannot sensibly ask clients to share detailed personal financial information.

However, these less intense programs can still ask clients to share some information. It is always a good idea to explain the “why” behind the data collection. This may include a desire to improve programs, help clients with goal setting, and get funding for future client services.

It may also be possible to collect data on anonymous surveys in some cases, to make clients more comfortable with sharing sensitive information. You could even ask clients to use a unique identifier or combination of information that only they know (like their full birthdate, middle name, or last four numbers of a Social Security number). This would allow you to compare clients to themselves over time without revealing their identity to you.

How many outcomes do I need to measure to improve my programs?

One. Any outcome measurement will improve your programs. Start small and measure one thing. Long term, you should aspire to measure 3-15 outcome indicators, depending on how complex your programs are.

"Piety is no substitute for technique."

ETIENNE GILSON

4. Collect your outcomes data (indicators).

Once you identify what to measure, the next step is to build a system to measure it. Simplicity is key. A simple system with clear instructions will allow you to collect accurate data regardless of who is doing the measuring.

HOW SHOULD YOU MEASURE?

The simplest way to measure is to roll up all your self-reported data in a single survey. This could be initially issued as a part of your program signup and then periodically from there. In most cases, the best practice is to have a staff member or regular volunteer fill out the form conversationally with the client. This turns the exercise into a relational bonding opportunity and ensures the client understands the instructions and purpose of the information collection. The staff member could be taking notes in the data collection form to speed up the process (digital system, spreadsheet, or digital form that populates to a spreadsheet). If the data entry is disrupting the relational connection, it could be inputted immediately following the conversation. Here is a [comprehensive self-assessment form](#), which you can modify to suit your needs.

Another option is a detailed evaluator survey, in which a staff member or volunteer is making the assessment himself rather than relying on client reported data. For instance, a program that attempts to empower people to improve their housing could send an evaluator to do a walkthrough with the resident/homeowner home before the intervention. The evaluator could have a 20-question survey with scores on each question ranging from 1-5, creating a maximum score of 100. This survey could be given before, after, and 12 months after the intervention. The final walkthrough is important if the program is attempting to empower the residents to maintain their own home. [See sample evaluator questions here](#).

For assessing knowledge before and after a class, a pre and post test is a good technique. [Click here for tips on designing a good test](#).

Be sure to test your surveys out on staff and more advanced or former clients before you begin using them. Get feedback on the quality, clarity, and emotional effect of the questions.

INTEGRATE MEASUREMENTS INTO CLIENT SUPPORT/CASE MANAGEMENT

As noted above, outcomes measurement is not only a way to improve your programs and explain their value to donors—good measurement also helps clients set and meet personal development goals. Data collection should be seamlessly integrated into your case management conversations.

Always keep your clients informed about what information you are collecting and why. Additionally, do not collect information that is likely to uncover what you are unequipped to address. There are scales and tests developed by psychologists to classify things like trauma, stress, and depression. It is unwise to ask clients these sorts of questions if you do not have the expertise to address the issues that arise or a referral system in place to ensure clients receive the help they need.

HOW OFTEN SHOULD YOU MEASURE?

You should always measure clients' situations as soon as possible after they join your program, and when they complete it (whether they graduate or not). Clients who drop out of a program may give you valuable information to improve the program, and they may have seen some improvement in their situation, even if they don't complete the program.

If a program is relatively short, such as a 12-lesson class, this before and after measurement may be all you need. If a program is longer, or more intensive (such as a 12-week residential therapy program) measurement in progress is a good practice.

Weekly or monthly measurement is a good practice when a client is pursuing rapidly evolving small goals that he or she is achieving in increments. Frequent touch points help maintain accountability and motivation. The intermediate measurements may be a shorter form of the total number of items you are measuring in the initial and final evaluations.

For long-term programs (12 months or more) quarterly or biannual measurements are a good time frame.

POST-PROGRAM FOLLOW-UP MEASUREMENT

Measuring at the completion of a program can be very difficult (say 3-12 months after program completion) and isn't necessary when you first begin measuring outcomes. However, in the long run, this kind of measurement offers valuable information about the program's durability. We strongly encourage every mature program to do at least some post-program follow up, even if only for a sample of clients on an occasional basis.

It is worth noting that the more time that has elapsed since program departure, the more difficult it is to connect a person's status to the program's effectiveness or ineffectiveness.

The biggest challenge is usually maintaining contact information and getting former clients to respond. The maintenance of contact information requires making sure you have redundant information when a client leaves a program. This includes multiple ways to contact the client (phone, email, social media accounts). If you're dealing with a transient population, consider collecting contact information for some more stable family members who are likely to keep in touch with the client.

To get clients to respond, the best way is usually to offer some incentive.

- One homeless and addiction recovery program offered free cosmetic dental work to clients who maintained their sobriety and stability for at least 3 months after the program completion. This incentive served both to help clients maintain their sobriety and created a touch point that would allow measurement of their post program success. By offering something valuable, they could also get feedback from nearly all clients. Even the ones who didn't come back to claim the reward were signaling that they had likely not maintained their sobriety.
- While the price point for that intervention was high, it doesn't have to be. You could offer people a gift card in gratitude for their time and achieve much of the same effect.
- A "moral incentive" can also be helpful. Make it clear to clients when they leave the program that you will be following up and that their information will help others. Whether they are happy about their situation at the time you reach out or not, they should know that their information will help improve and fund the program for people who come after them.

PROTECTING YOUR CLIENT'S INFORMATION

Data protection is an important part of protecting your clients. Avoid collecting sensitive information that is not critical. Don't use a Social Security number for administrative purposes if you could use name and birth year or some other less personal piece of information. If you need to collect sensitive information, destroy what you do not need to keep on file. For instance, if you allow clients to pay program fees by credit card, don't keep their card information after they leave the program.

For data you do retain, keep client data in a separate space (digital or physical) from routine organization documents. This could be an online tracking system (like CharityTracker), a separate digital file sharing platform, a shared drive folder with limited access, or in a locked physical filing cabinet.

Let clients know how you are protecting their information and privacy. It will help build trust.

It is a good practice to get written consent before you share your client's stories with supporters, particularly if you are using their name, image, or other details that could clearly identify them.

5. Compile and analyze your measurements.

After you collect the data, you'll need to compile it to see averages, trends, and changes over time.

WAYS TO COMPILE THE DATA

Use a specialized software tool.

There are specialized software tools like Charity Tracker Plus and Mission Tracker that allow you to track your client data and outcomes in a digital system. These can be great solutions, but for many smaller organizations, they may be too expensive or cumbersome to manage. Even if you use one of these systems, you still need to go through the process above to determine what your outcomes are and you need a basic understanding of how to collect reliable information.

Use a spreadsheet.

In most cases, smaller organizations will be collecting data into spreadsheets and then compiling multiple clients into a single sheet periodically. In order to make your spreadsheet easy to draw insights from, it needs to be formatted sensibly. Here are some best practices for formatting:

Place individual clients in rows, and headers in columns. This makes it easier to use automatic filters to sort clients by results (e.g. separate those who found a job from those who have not)

Put as much data as possible into numerical formats. Yes/No options can be turned into 1's and 0's. This allows easy counting and averaging of the number of yeses, using built-in count and average features. Details like "length of time in program" should always be in a standard format like number of days or weeks, rather than writing "two months" in the field. This allows you to sort clients by length of time in the program and see how the impact of your program changes clients over time.

Standardize text inputs when possible. Some fields, such as name, are unique for every client. However, for many other types of text information, such as "living situation," there may be a limited range of options (e.g. rent, own, living with family, living with friends, homeless, unknown). In order to simplify filtering and counting, use a dropdown box to

standardize the options for the field. That way, even if you have multiple people entering data, everyone will choose from the same options, which will simplify analysis.

Use “data validation” tools to standardize inputs. Most spreadsheet programs have a set of built-in tools for “data validation.” These tools allow you to customize portions of the spreadsheet such that only certain types of data are allowed, such as numbers, text, or dropdown boxes. You can also lock certain cells so that no one can accidentally delete formulas, headers, and other parts of the spreadsheet that are not intended to be edited.

Use formulas to auto-calculate as much as possible. You can create formulas to automatically calculate averages, changes, sums, and many other useful features. Set them up in the rows above the headers and it won't prevent you from using filters to reshuffle clients.

[Here's a spreadsheet](#) that incorporates many of these best practices which you can use as a template for your work.

HOW TO ANALYZE THE DATA

Turning Numbers into Insights

Now that you have the information in one place, you need to turn the data into usable and shareable information. Here are some things you can look for:

DO YOUR CLIENTS IMPROVE ON AVERAGE?

The best place to start is to look for change in your clients situations. Start with the average. What is the average credit score on program entry? What is the average score at the subsequent measurement? Do this for each indicator you measure.

What you could learn from this: Average change is the most basic measure of a program's success. Are most people better off when they leave than when they came? If things are not changing on average, you need to rethink your approach. There may be some external shock (like a recession) that could cause this temporarily, but if you rule that out, then it means the program needs work. A good way to start might be to see if there are any clients for whom it seems to be working. Ask yourself what's different about them. Are they more ready for change? You could use more self-selection in your screen process to ensure you're working with those most likely to benefit from your program. Do they have some set of skills or knowledge that is helping them? Perhaps you could provide that to the other clients as well.

WHAT FRACTION OF CLIENTS ARE MOVING IN THE RIGHT DIRECTION?

It is possible for clients to improve “on average” even if 90% are stagnant as long as the remaining 10% made a sufficiently large change. To avoid this pitfall, sort clients by whether they are making positive change. You can use formulas to build a column that sums the

difference of before and after measurements (e.g. Pre test 82, Post Test 70, change = -12). If you sort clients by that column, you'll be able to see what fraction are moving in the right direction.

What you could learn from this: In most cases, the majority of your clients should be making progress. If they are not, go looking for root causes (which could also include measurement errors).

DOES THE CHANGE INCREASE OVER TIME?

If you have an intervention that lasts a standard amount of time (like a six-week class), then you only need to measure before and after. However, many interventions, like food co-ops, transitional housing, or shelters see clients for varying amounts of time. In this case you'll want to sort clients into groups based on how long they have stayed with your program. Two or three groups is usually sufficient (e.g. less than 3 months, 3-9 months, more than 9 months). Ideally, you should see that positive change increases as the treatment increases. This may be reflected in better averages in later groups, or in increased change from their own baselines.

What you could learn from this: You may find that people experience rapid growth from your program, but then tend to plateau at a certain point. This could inspire you to make it more difficult for people to stay past a certain duration in order to free up the program for people who will benefit more. Or you may find they are hitting some common obstacle which is causing the plateau, that you may be able to address. You may find that people who stay in your program for only a short amount of time experience limited change. You could use this information to structure your program to get people to commit for a minimum amount of time after a trial period, or if that fails, you could provide an incentive for people to make it to some point.

COMPARE YOUR PROGRAM TO BENCHMARKS.

You may be disappointed that only 40% of your clients graduate from your rehab program. However, it's important to know that the average graduation rate from a 3-6 month rehab program is only [20-30%](#). This benchmark is an important measure of your success. If you can't find a good scientific benchmark for your program, contact similar programs (such as fellow [members](#) of the True Charity Network) to find out their success rates. The conversations that follow may help improve both programs.

MAKE OBSERVATIONS ABOUT INDIVIDUALS.

In addition to insights that will affect your programs, you can also use data to serve your clients directly. You may observe individual clients are regressing or plateauing. It is often the case that your case management conversations will have already revealed this information, but clients are sometimes reluctant to acknowledge that they are struggling.

You can use the measurable evidence found in your data to provide simple proof that they need to do something different.

Pitfalls To Avoid

DON'T BE DECEIVED BY OUTLIERS.

Maybe a client inherited \$200k and paid off her mortgage and all of her debt. It's not reasonable to allow that fact to shift your average debt reduction by 50% for dozens of clients. Make sure outliers aren't having an outsized impact on your averages.

DON'T CONFUSE REAL TRENDS WITH RANDOM FLUCTUATION.

It's natural to want to claim credit for any positive change (and claim that any negative movement is unrelated to your program). However, in order to be as honest with yourself and stakeholders as possible, you will have to do your best to distinguish the "signal" of real change from the "noise" of random variation due to clients' life circumstances unrelated to your program. In general, you want to see a larger sample of clients, and a larger magnitude of change, to be confident that the change resulted from your program. This chart is a general guideline in how to increase your confidence in your results.



"It's no longer good enough to make the case that we're addressing real needs. We need to prove that we're making a real difference."

MARIO MARINO, AUTHOR OF LEAP OF REASON

Table 4: Differentiating Real Change from Random Fluctuation

	Small Population	Large Population
Small Change in Indicators	Likely random variation (e.g., 5 clients, average decrease in debt of 5%)	Likely real change (e.g., 50 clients, average decrease in debt of 5%)
Large Change in Indicators	Likely real change (e.g., 5 clients, average decrease in debt of 50%)	Almost definitely real change (e.g., 50 clients, average decrease in debt of 50%)

DON'T CONFUSE CORRELATION WITH CAUSATION.

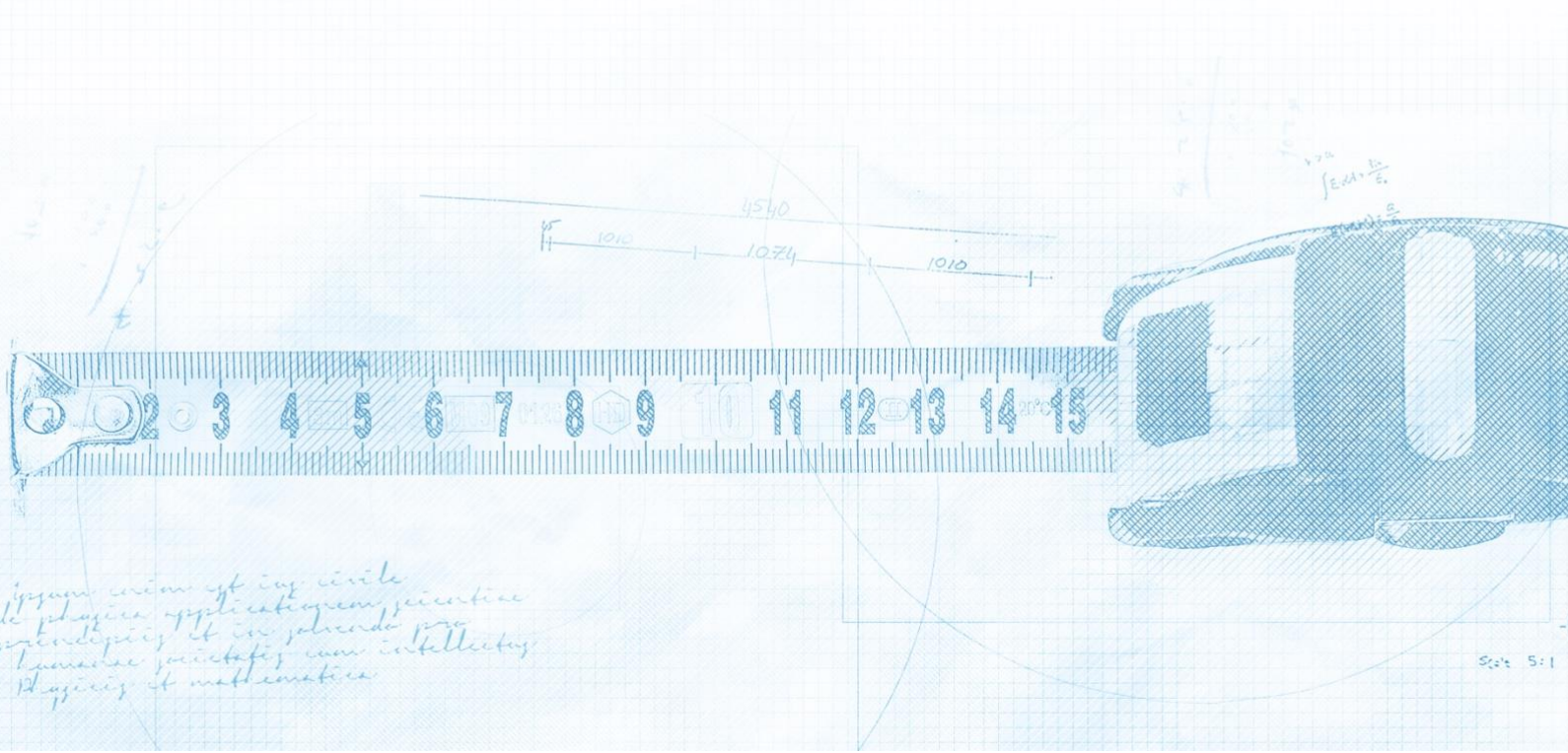
Even when you are confident that a real change happened in the outcomes of your clients, it doesn't necessarily follow that your program caused the change. It could be that some other major non-programmatic factor caused the shift. For instance, maybe you measured net worth in the quarter before and after clients received their annual tax refunds. Short of a scientific randomized trial, it may not be possible to eliminate this risk entirely. However, if you make a good faith effort to eliminate other major factors outside of your program, it is

WATCH OUT FOR CHANGING EXPECTATIONS.

When you measure indicators on subjective scales, people's expectations affect their answers. For example, if you ask a man how financially secure he feels on a scale of 1-10 and he works a minimum wage job, what do you suppose he will answer? It depends on his expectations. If he is a college grad, now working for minimum wage, he will probably give a very low answer (such as a 2 or 3). If he is a recent immigrant from a developing country, he will likely give a much higher answer (9 or 10).

If the immigrant is in your program, he may realize that he could do much better than minimum wage. When he leaves your program, perhaps he is earning \$15 an hour, but if you ask him how he perceives his financial security on a scale of 1-10, he may now answer a bit lower than when he entered the program, despite his objective higher wages. This doesn't mean he is worse off, or even unhappy with his progress, it only means he has expanded his sense of what is possible.

The same effect can occur in any area. As people grow in relationships and capacity, their expectations often expand as well. If you see people's subjective ratings deteriorating, especially in the first few months after they enter your program, investigate to see if their expectations are changing.



What to Do With Your Outcomes

Troubleshoot to Improve Your Outcomes

A common objection to measuring outcomes (rarely stated) is that you may not like what you discover. In many cases, the process of outcome development itself will result in the recognition that new programs are needed, or that existing ones may not be suitable. Even when programs are sensible, you still may find that your program underperforms benchmarks. The previous section identifies how your analysis may be able to uncover root problems (perhaps your program admits a lot of people with lower-than-average commitment or has an exceptionally difficult element too early in the program for clients to be ready to overcome it).

DIVE DEEPER INTO YOUR CURRENT SITUATION

In addition to the basic insights you can gather from the preceding section, here are some other techniques you can use to improve your programs:

What can you learn from your bright spots?

Look at your top 10% or so of clients in terms of their change. Do they have any common characteristics that contribute to their success? Could you screen for characteristics like these in your application process? It may be worth speaking to a few of them to learn more about what they attribute their success to. Did they have an above-average mentor or read a book that kept them inspired? Is there anything you could do to pass their techniques on to other clients?

What went wrong with your dropouts?

Many clients do not live up to your hopes for them. See what traits they have in common. Do they tend to drop out at a certain point in the program? You may be able to strengthen their support or motivation at that critical point. Do they have traits that you could screen out at application in favor of clients more likely to benefit from your programs?

Look for relationships between indicators.

Look at variables that seem to be logically related and use sorting capabilities in your spreadsheet or data management tool to see if the results in one category are related to the other. For instance, a homeless shelter tracks whether residents comply with their recommended 75% savings program. By comparing compliance status against the type of housing people moved into upon departure, we learn that clients who comply have a 35% rate of leaving with independent housing, clients who do not have a 6% rate. Knowing the importance of this program element may cause them to encourage it more deliberately or experiment with making it mandatory.

Research other organizations.

Check out the [models of charity archive](#) to find programs working with similar clients - reach out to them through the [directory](#).

ORGANIZE YOUR INSIGHTS

With the information generated by this analysis, you may already have clear ideas of program elements to modify. If you don't, consider these techniques to uncover and organize problems and solutions. These exercises work well for an individual or a team. You will naturally need client input to complete these assessments. You can either create a simple conversational questionnaire for a sample of clients or you could just bring a few clients into a session with staff members and collectively attempt to identify the causes. Consider the following approaches and choose the one that works best for you:

Force Field Analysis

This is a simple diagram listing the forces that would compel a client in the right or wrong direction. Identifying these forces is the first step to generate ways to boost forces for change and minimize those against it.

Example

Finance class with outcomes of debt reduction and savings increase

Forces for Change

- Learners are tired of living paycheck to paycheck
- Learners are motivated by program graduates who show change is possible
- The curriculum is full of clear guidance for effective money management techniques

Forces Against Change

- Learners' friends and family ask them for money when they get ahead
- Learners' habits prioritize instant gratification
- Learners' high-fee bank accounts drain their savings
- Government benefit asset thresholds discourage learners from saving
- Family situations (single parenting and/or child support) make it difficult to save
- Personal emergencies derail emergency funds and force high-interest debt

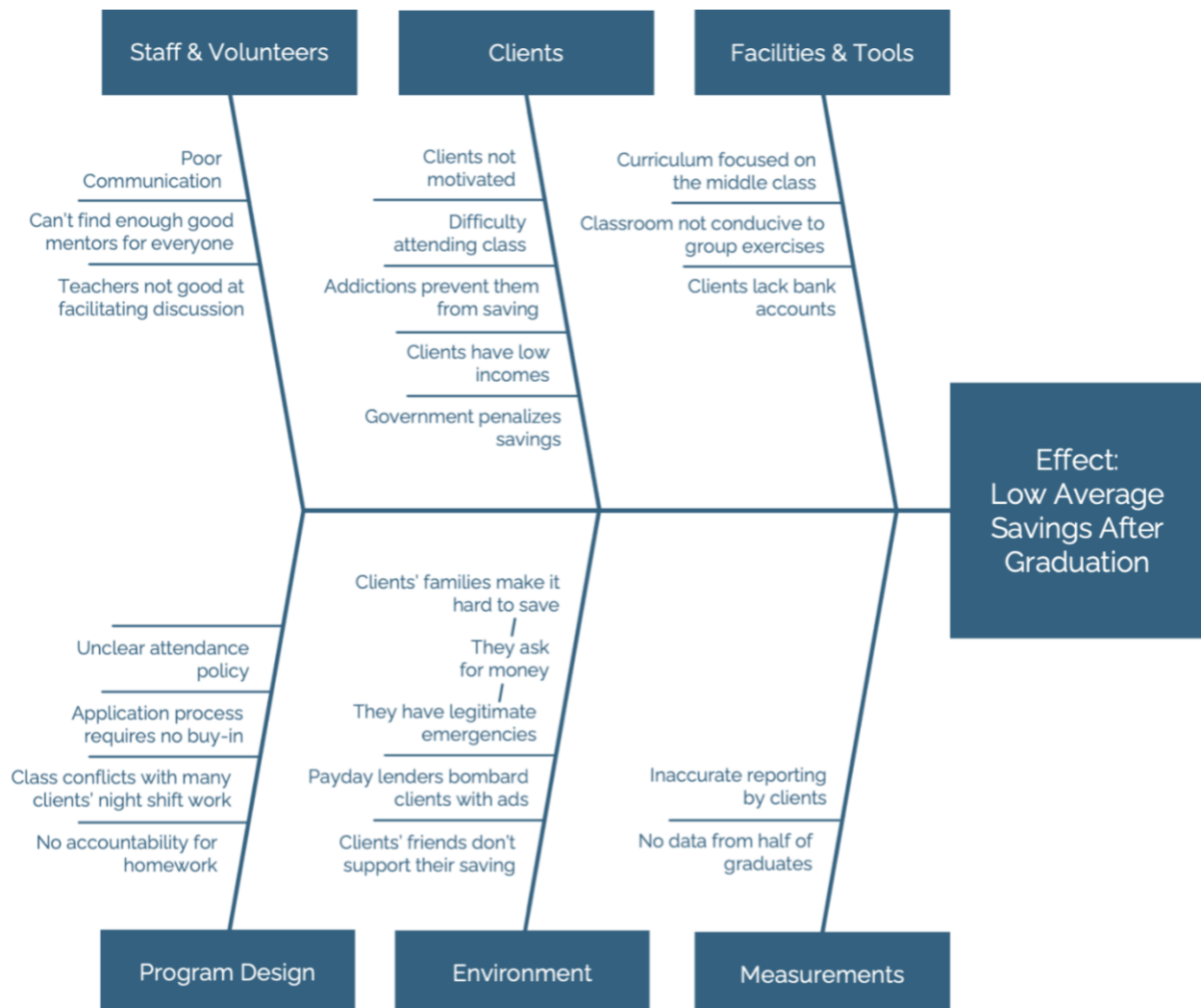
In this example, detailed attempts to address the forces against change may generate solutions like Individual Development Accounts (exempt from government benefit asset limits), partnerships with better banks willing to offer lower fees, alternative financial models like savings groups. You could also add forces for change, like a matched savings program, or finding partner employers willing to give learners a chance to earn additional part-time income.

Cause and Effect Diagram

This diagram technique tries to identify multiple causes for a single effect, such as low graduation rates. It is essentially a more detailed dive into one side of the force field diagram. Start with the following six high-level causes, and then add sub-causes to each one.

- 1) Staff and Volunteers
- 2) Client
- 3) Facilities and Tools
- 4) Program Design
- 5) Environment
- 6) Measurement

Example: Finance Class



Note that it is not important that you assign each cause to the "right" category, it is only important that you identify all the possible causes, so that you can then identify ones which you may be able to address.

Brainstorming with Affinity Diagrams

You can use brainstorming to either identify possible causes (in conjunction with the force field analysis of cause and effect diagram) or to identify possible solutions to those problems. In either case, have each member of the team independently generate multiple ideas on paper, then have everyone group and discuss their ideas. Group like ideas and duplicates into categories. You can do this physically if people write their ideas on sticky notes or strips of paper. These categories form the basis of your solutions.

EXPERIMENT AND MONITOR.

Once you implement your new solutions, you can use your outcomes to monitor whether they have the desired effect. For example, if you add Individual Development Accounts (IDAs) to your financial program, you should monitor to see that clients actually use them, and to see if the clients that use them demonstrate better savings results than those who do not. Finally, ensure that this is increasing your average outcomes. Otherwise, it may be the case that your top 10% of learners adopted the IDAs, but didn't save any more than they would have otherwise.

EXECUTE WITH EXCELLENCE.

Some solutions may not involve a clean program change, like a shift in curriculum. You may need to improve elements that require continuous monitoring. For example, you may need to ensure that classes are more interactive or that volunteers are more relational. For these kinds of improvements, the four disciplines of execution (4DX) may prove useful ([learn more about them here](#)).

One good way to ensure processes stay on track in medium to large organizations is to institute an "operational audit." This is a systematic review of effectiveness, efficiency and economy of operation in an organization. These are usually performed on an infrequent basis and require an external auditor who is a specialist. If this seems too complex, you can do "operational reviews" internally and accomplish some of the same objectives ([explained in depth here](#)).



Share your Success

HOW TO PRESENT YOUR OUTCOMES TO SUPPORTERS

Now that you have compiled your data, it's time to share it with your supporters. For most donors, you will probably only share the top 5-10 pieces of information that best represent your client's progress. This information could go on a social media info-graphic, on a printed card or brochure, or in a presentation ([see examples here](#)). Here are a few best practices of what to share:

Report the sum of the change.

For some outcome indicators, you may report the sum of the change, such as the sum of GED/HISET tests passed or family reunifications. If your clients became independent from certain government benefits (and started paying taxes), you may be able to calculate an impressive sum of taxpayer dollars saved.

Report the average change.

For some outcomes, you may want to report the average change, such as the average savings increase or average credit score increase.

Report the number of clients making positive progress.

Different clients may be progressing in different areas. Report the number of clients making progress over a set period of time to capture all of their improvements.

Report performance relative to benchmarks.

As noted previously, your results may be discouraging in isolation, but when weighed against benchmarks, you may be at or above average. Be sure to provide that context for donors.

Share things you want to improve.

You don't need to trot out every shortcoming of your program, especially if you're working on fixing it. However, donors value transparency, and it can be a good idea to use your measurements to highlight an item or two that you want to improve and explain how a donor's support will help you do it (by adding a program or resource).

Be honest.

It is possible to use out of context statistics to trick people. Don't do it. Start by being honest with yourself, and if you think your programs are making a difference, then make the case (if

you think your programs aren't making a difference—change them!). If your graduation rate isn't great, you can work on improving that, while still reporting that those who do graduate are succeeding.

Share stories!

Outcome measurements don't replace stories, they amplify them. Stories are powerful, but don't prove much on their own. Even an organization with a 98% failure rate will probably be able to come up with a couple of positive stories for a fundraising dinner. By all means, tell your best stories, and then your outcome measurements show that your stories are representative of most of your clients. Your data proves your anecdotes aren't flukes.

Note: A common issue that arises when reporting change is the distinction between "percent improvement" and "percentage point improvement." For instance, if 3 in 10 people budget before a finance class (30%) and 9 in 10 (90%) of them budget after going through a class, there are two correct ways to report the change. It is correct to say that *budgeting increased by 200%*, because there are six new budgeters, who represent two times as many as the initial 3. It is also correct to say there was a *60-percentage point improvement* in budgeting, because budgeting rates shifted from 30% to 90%. It is incorrect to say there was a 60 percent improvement in budgeting, because 60% of the three initial budgeters is 1.8, which would imply an improvement to 4.8 total budgeters, not 9. To eliminate confusion, it is generally best to simply show the before and after percentages.

SHARE YOUR SUCCESSES WITH STAFF AND CLIENTS

In our desire to please donors, it's easy to overlook the importance of our staff and volunteers. It is critical that they understand how data is informing programs and demonstrating success. It is especially encouraging to the people who go to great lengths to collect or compile the information to know that it is being put to great use.

As mentioned previously, it is also important to use data in case management. Not only can it be helpful in tracking individual progress, but it can also be worthwhile to let clients know where they stand relative to the average of their peers. This must be done selectively, when you think it would inspire and not discourage a client. It would be inappropriate to post a ranking of clients for all to see, but it may motivate some to know that they are performing well or even to know that they are lagging behind. It can also be used to encourage them to take specific recommended actions that are associated with better outcomes in their peers.

Refine As You Go

Outcomes measurement is a long-term investment. Your initial start may be rough, but you can refine everything over time.

Every time you do your data analysis, ask yourself what information you are not using. You may find that some indicators are so stable that they are uninteresting or so wildly variable that you clearly cannot control them. In either case, it is fine to stop measuring things when you know you are not going to do anything with the information you gather. Conversely, as you identify problems and bright spots, you will also likely find other outcomes that you would like to start measuring.

