impactmanagementproject.com

Using self-reported data for impact measurement

How to use stakeholder surveys to improve impact performance

IMPACT MANAGEMENT PROJECT

60_decibels

keystone

ON
OMIDYAR NETWORK

Contents

About the Impact Management Project

The Impact Management Project (IMP) is a forum for building global consensus on how to measure and manage impact.

We convene a Practitioner Community of over 2,000 organisations to debate and find consensus on technical topics, as well as share best practices.

We also facilitate the IMP Structured Network – an unprecedented collaboration of organisations that, through their specific and complementary expertise, are coordinating efforts to provide complete standards for impact measurement and management.

This pilot was sponsored by Omidyar Network

Omidyar Network is a philanthropic investment firm that invests in entrepreneurs and their visionary ideas to create opportunities for people to improve their own lives, their communities, and the world around them.

Since 2004, Omidyar Network has committed over $1 billion to for-profit companies and nonprofit organizations that foster economic advancement and create opportunities for individuals across six key areas: Digital Identity, Education, Emerging Tech, Financial Inclusion, Governance & Citizenship Engagement, and Property Rights.

Foreword by

Omidyar Network

Over the last decade, the potential for all enterprises to measure and manage their positive and negative impact on society and the environment has become more widely understood. However, there is still far too much variation in the way different funders and enterprises think about and measure impact. That makes it incredibly difficult for different organisations to work together to create lasting social and environmental impact.

The Impact Management Project has made great progress in the past three years by bringing different communities together to agree on a common definition of impact, and the dimensions of performance that matter for impact measurement, management and reporting across the value chain.

But there is still much work to be done on how impact measurement can be carried out in a cost-effective, accurate and useful way that drives impact management decision-making. For example, impact data is often either missing altogether, pulled together from third party research (of varying relevance), or based on poor proxies pulled from operational financial data such as sales figures. Unreliable impact data is not conducive to good impact management.

There are a number of challenges to overcome to ensure all enterprises and investors measure and manage impact consistently and effectively.

Three key issues include:

- Misapplication of measurement methodologies;
- Insufficient measurement approaches that are cost-effective and easily repeatable; and
- A lack of common metrics, which would allow comparability and benchmarking over time.

We believe that using standardised stakeholder surveys designed to collect rapid data on impact performance - across each dimension of impact identified through the IMP - is part of the solution to these challenges.

At Omidyar Network, we're committed to impact investing and its role in supporting life-changing organisations around the world. Founded by philanthropists Pam and Pierre Omidyar, the founder of eBay, our firm has a deep history in pioneering approaches to invest and scale innovative organisations that catalyse economic and social change, and at the heart of our work is a commitment to learning and strengthening impact.

By supporting this pilot project, we hope to provide practitioners with easy-to-use survey tools that aid measurement of the five dimensions of impact and make it easier to infuse meaningful impact data into business decision-making.

This report is a first, critical step in helping funders and enterprises design and implement a survey that can provide meaningful impact data across all five dimensions of impact. This report includes a list of tried-and-tested survey questions for each dimension of impact, alongside practical insights to guide implementation and help organisations improve their own impact measurement and management practices.

We look forward to the debate that this report may spark within and beyond the impact investing community. Together, we can help impact measurement reach its full potential as a rigorous tool that enables greater positive social and environmental impact for generations to come.

Jessica Kiessel Director of Learning and Impact, Omidyar Network

Executive Summary

What we did

Through working with groups of enterprises and investors, the Impact Management Project was able to specifically identify what categories of data are needed to understand performance against each dimension of impact (see Chapter 1, p.5).

In partnership with 60 Decibels, Keystone Accountability and Social Value International, we then set out to design simple, repeatable customer and employee surveys which would facilitate the easy collection of impact data across all data categories for each dimension of impact (see Chapter 2, p.10).

60 Decibels and Keystone Accountability then pilot-tested these surveys with portfolio companies of Acumen and Bridges Fund Management, in order to share real examples and surface lessons and guidance relevant for any organisation seeking to use surveys to measure and understand their social and environmental impact (see Chapter 3, p.14).

What we learned

1. Surveys are essential for gathering data on the What dimension of impact, as these data categories require the affected stakeholders to report back on which outcomes they are experiencing, and which are most important to them.
2. An easy to implement, generic survey can give a first-impression of impact across most of the other dimensions of impact and is a great way to begin impact management.
3. A full understanding of the How Much and Contribution dimensions of impact will often require other impact measurement approaches.

How to use this report:

New to the IMP?

For an overview of the five dimensions see p.6 or visit www.impactmanagementproject.com

Want to see examples of self-reported data being used alongside non-self-reported data?

See Appendix 1, p.46-47

Want to find information about questions for specific dimensions?

Customer survey:

p.16 What outcomes are occurring?

p.20 Who experiences it?

p.24 How much change occurs (scale, depth and duration)?

p.32 What contribution does the enterprise make to what would likely have occurred anyway?

p.35 What is the risk that the resulting impact will be different than expected?

Employee survey: p.41-42

Introduction

Consulting stakeholders is the foundation of impact measurement and management

The purpose of impact management

Different enterprises have different impact and financial goals. Some seek to mitigate unwanted harm, while others will be able to go a step further and also try to generate positive change.

Regardless of goals, the central purpose of impact management is to reduce negative impact and improve positive impact. Any enterprise that hopes to do this must begin by understanding the impact it currently has on people and the planet. This requires impact measurement. Once it has gathered and analysed this information, the enterprise can then set targets and seek to improve its impact performance over time.

What kind of measurement is needed

Measurement can be approached in stages. Enterprises might start by understanding who their customers or wider stakeholders are, by collecting data on their demographics and seeking feedback on their wants and needs (What). An enterprise might then seek more specific feedback from their stakeholders about what other kinds of impact, positive and negative, they are experiencing when engaging with the enterprise – either through using the product, service or through the operations of the enterprise as an employee, community member or supplier.

To manage these impacts, the enterprise will then need data on which impacts are most important (What) to the stakeholder affected, as well as the degree of change (How Much) that has occurred for each.

Many enterprises start this process by looking to third-party data on the impact of similar enterprise models. However, unless the type of impact the enterprise is seeking to deliver is unaffected by context (which is occasionally if rarely the case) this will only ever provide a proxy for impact (and may or may not be a good proxy). Therefore, given third-party data will often not be applicable, the greatest accuracy will always come from the enterprise gathering new data about its own impact on its stakeholders.

There are several approaches that can be used to collect new data. Each provides varying degrees of certainty, and some are more appropriate and effective in different contexts than others. This paper looks specifically at the role of surveying to collect impact data across the five dimensions of impact identified through the Impact Management Project.

For the purposes of this paper, we have referred to the data collected through surveying as self-reported data, and have sought to illustrate how this data is particularly valuable when beginning to measure the impact of an enterprise. Self-reported data will also be shown alongside non-self-reported data to illustrate how these approaches are complementary; in fact, they can often be used interchangeably across the dimensions of impact (see Appendix 1 for an example).

It's important to recognise that none of these data types are necessarily better or more rigorous than any other. Each has relative strengths and weaknesses. The best option to choose/use will depend on the type of impact or business model in question (as illustrated by Figure 1 below).

Non-self-reported data is used here as a catch-all term for data that indicates achievement of the outcome that has not been verbally reported by the stakeholders themselves. This data can vary in terms of their level of rigour from a weak proxy (e.g. people attending a course to quit smoking as a proxy for their reduction in smoking), to more evidenced proxy data (e.g. minimum hours of exercise of school children per week) or evidenced outcome data (e.g. measured blood sugar level of diabetics).

Self-reported data is collected directly from people who interact with the enterprise (e.g. employees, customers, suppliers etc). Self-reported data can be either objective or subjective, as illustrated in Figure 1.

FIGURE 1 Types of Data

Subjective self-reported data	Objective self-reported data	Objective non-self-reported data
"I know more than I did before the course"	"I was offered a job shortly after completing the course"	Student completed and passed the course with a mark of 80%
"I am coughing less this month"	90% of customers substituted the kerosene lantern for the solar lamp	500 solar lamps purchased
"I feel better this year compared to last year"	"I went to a doctor three times last year, compared to 15 times the year before"	Patient blood pressure: **2017:** 150/90 **2018:** 125/80

Who this report is for:

In the last two years, the Impact Management Project has consulted with over 2000 practitioners to build consensus on norms of impact measurement and management. Through all of these conversations, the importance of the stakeholder voice has been a recurring theme.

Organisations of many different kinds, and with a variety of different perspectives, all recognised the importance and value of understanding what impact is occurring from the perspective of those affected. And yet few felt that they had adequate access to resources on how to collect and use this data – and in particular, what questions they should ask.

This report, and the pilot it describes, seek to provide guidance on using feedback for any organisation looking to better understand and manage their impact on people and planet.

Any questions, contact:

The IMP: team@impactmanagementproject.com or visit www.impactmanagementproject.com

60 Decibels: info@60decibels.com or visit www.60decibels.com

Keystone Accountability: info@keystoneaccountability.org or visit www.keystoneaccountability.org

Designing a survey along the five dimensions of impact

To understand any impact on people or the planet, data is needed across five dimensions: What, Who, How Much, Contribution and Risk.

Context: the five dimensions of impact

Through the IMP, more than 2,000 practitioners have reached consensus that any type of impact experienced by people or the planet – intended and unintended, positive and negative – needs to be understood across five dimensions (as shown in Table 1 below).

TABLE 1 The five dimensions of impact

Dimension	Questions on each dimension
What	What outcome(s) do business activities drive? How important are these outcomes to the people (or planet) experiencing them?
Who	Who experiences the outcome? How underserved are the stakeholders in relation to the outcome?
How Much	How much of the outcome occurs across scale, depth, and duration?
Contribution	What is the enterprise's contribution to what would likely happen anyway?
Risk	What is the risk to people and planet that impact does not occur as expected?

The IMP reached global consensus that impact can be deconstructed into five dimensions: What, Who, How Much, Contribution and Risk.

The IMP's starting point for developing these five dimensions was to work with teams at Impactt and Social Spider, who delivered focus groups in Myanmar, Kenya, Belgium, United Kingdom and Bolivia[1]. The aim of the focus groups was to help the IMP understand what people's expectations are when buying from or engaging with products, services or opportunities that they want, in order to understand how people think about themselves and their goals, as stakeholders of an enterprise.

By involving the whole value chain — such as policymakers in Argentina, fund managers in East Africa, and asset owners in the UK — to test these five dimensions, it became clear that understanding a stakeholder's experience across all five dimensions is necessary for comparing one impact with another. If data on one dimension is missing, performance cannot be appropriately understood. For example, the impact of a job created for a middle-income, well-educated person is not the same as — and therefore should not be compared with — a job created for a long-term unemployed person.

Collecting data for each dimension therefore enables enterprises, and their investors, to assess and manage their impact in a consistent and comparable manner, and then set goals against one or more dimensions to improve their impact.

The indicator best suited to measuring impact will vary considerably in different contexts; so, the IMP has focused on providing guidance on the **15 categories of data** that will enable the appropriate analysis for each dimension, as illustrated by Figure 2 overleaf.

[1] Read about the insights gathered in this process here

The impact data categories

The impact data categories specify the type of data that enterprises can use to set goals and understand performance for each dimension.[2]

FIGURE 2 The 15 impact data categories

Dimension	Impact data category	Description	Is the enterprise…	
What What outcome(s) do business activities drive? How important are these to the people (or planet) experiencing them?	1 Outcome level	The outcome level experienced by the stakeholder when engaging with the enterprise. The outcome can be positive or negative, intended or unintended.	generating a negative outcome?	generating a positive outcome?
	2 Threshold for positive outcome	The level of outcome that the stakeholder considers to be positive or 'good enough'. The threshold can be a nationally- or internationally-agreed standard.	delivering a level of outcome above or below a nationally- or internationally-agreed performance standard?	
	3 Importance of outcome to stakeholder	Stakeholders' view of whether the outcome they experience is important	generating unimportant outcomes?	generating important outcomes?
	4 Sustainable Development Goals or other globally-recognised goals	The globally-recognised goal(s) that the outcome relates to		
Who Who experiences the outcome? How underserved are the stakeholders in relation to the outcome?	5 Stakeholder	The type of stakeholder experiencing the outcome		
	6 Geographical boundary	The geographical location where the stakeholder experiences the social and/or environmental outcome.		
	7 Baseline outcome level	The level of outcome experienced by the stakeholder prior to engaging with the enterprise	reaching well-served populations?	reaching underserved populations?
	8 Stakeholder characteristics	Socio-demographics and behavioural characteristics of the stakeholder to enable segmentation		
How much How much of the outcome occurs across scale, depth, and duration?	9 Scale	The number of individuals experiencing the outcome	generating the outcome for few?	generating the outcome for many?
	10 Depth	The degree of change experienced by the stakeholder.	delivering a small degree of change towards the outcome?	delivering a large degree of change towards the outcome?
	11 Duration	The time period for which the stakeholder experiences the outcome	generating short-term change?	generating long-term change?
Contribution What is the enterprise's contribution to what would likely happen anyway?	12 Depth counterfactual	The estimated degree of change that would occur anyway for the stakeholder	contributing marginally or not at all, relative to what would have happened anyway?	contributing significantly, relative to what would have happened anyway?
	13 Duration counterfactual	The estimated time period that the outcome would last for anyway		
Risk What is the risk to people and planet that impact does not occur as expected?	14 Impact Risk type	The type of risk that may undermine the delivery of the outcome	taking a low level of impact risk?	taking a high level of impact risk?
	15 Impact Risk level	The level of the risk specified in risk type (e.g. High, Medium, Low)		

See Appendix 1 for an example of how enterprise data can be mapped to these categories in practice. In the example, data for just one of the enterprise's effects – weight loss for customers – has been mapped. It seeks to illustrate how both self-reported and non-self-reported data can produce different, and often complementary, insights which leads to a fuller understanding of the impact occurring.

[2] For additional guidance on each one of the categories, please refer to the What is Impact section of the IMP website

Pilot-testing survey questions for each dimension

Our objectives:

1. **Co-create questions across the five dimensions to enable enterprises to begin to assess their impact, and to open up the possibility of benchmarking across standardised questions.**
2. **Gather insight on what questions work, as well as the limitations to the standardised questions in terms of their ability to capture the right data for each of the dimensions.**

Step 1: Design survey questions for each of the five dimensions
Our goal was to create a generic set of questions that were agnostic of the outcome, sector or demographic. We also wanted to learn much as possible about impact for each dimension with as few questions as possible.

Teams from 60 Decibels, Keystone, Social Value International and the Impact Management Project pooled existing, evidenced survey questions to identify which were most likely to gather high quality feedback about each dimension of impact. Where questions did not exist, we crafted new questions and undertook initial testing.

The challenge was to be specific and generic at the same time: specific enough so respondents could easily understand what we were asking, and yet generic enough so the survey could apply across different scenarios and sectors.

Step 2: Test surveys through a series of pilots in different geographies using different collection mechanisms
The surveys were piloted using different data collection tools (phone calls, in-person surveys and online surveys). Occasionally the questions had to be phrased differently to perform most effectively. But all surveys kept the same common rubric to ensure data is captured for each of the five dimensions of impact was as comparable as possible.

The enterprises surveyed are portfolio companies of Bridges Fund Management, a UK-based impact investment fund manager (surveys carried out by Keystone), and of Acumen, a global non-profit (surveys carried out by 60 Decibels).

The Bridges and Acumen teams facilitated access to the enterprises. Keystone and 60 Decibels then engaged directly with the enterprises to run the surveys. Findings were shared with both the enterprises and the investor so that both parties could discuss and agree on how the data should be used to improve performance.

Step 3: Build on existing expertise
The team shared early findings and sought feedback from peers through the Feedback Labs conference 'Feedback Summit', and by comparing findings and approaches with the DFID Impact Fund's deep dives, also conducted by 60 Decibels[3].

This work also explicitly sought to build on the consensus-building work on this topic, carried out by a working group of practitioners through the World Economic Forum's initiative on 'Mainstreaming Impact Investment' in 2017.

---

[3] Webinar recordings of these findings co-hosted by DFID and IMP can be found here and here.

The pilot enterprises

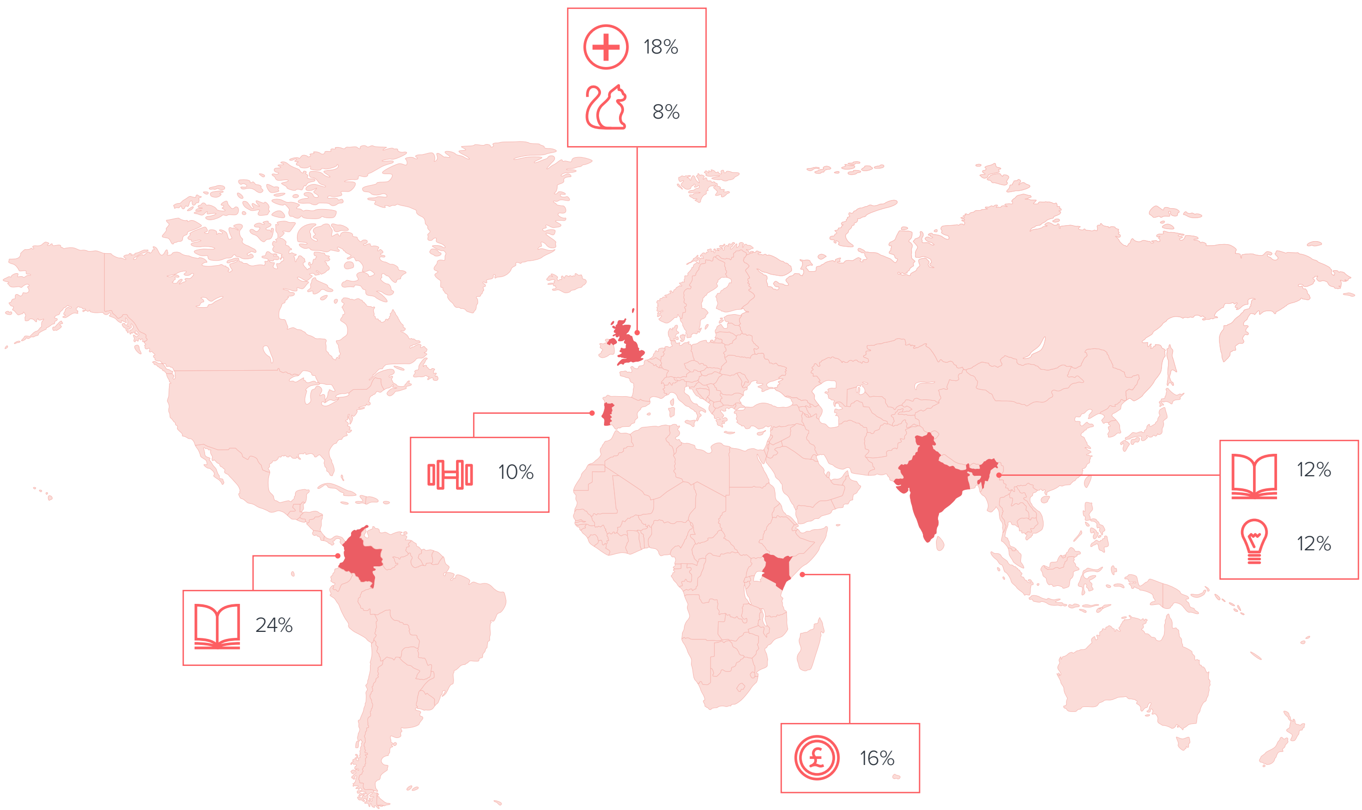
60 Decibels surveyed four portfolio companies of Acumen:

- **Edupol** provides technical and vocational training to BoP population living in Colombia's remote municipalities by serving as a bridge between higher education institutions and students.
- **EduBridge** offers customised job training and placement program catering to students with high-school diplomas or dropouts from urban or peri-urban training centers in India.
- **Juhudi Kilimo** provides microfinance loan products that allow Kenyan smallholder farmers to access high-quality agricultural assets that enhance the productivity of their farms.
- **Husk** takes agricultural waste, rice husks otherwise left to rot, and converts it into gas that powers an off-the-shelf turbine to generate electricity for Indian villages.

Keystone Accountability surveyed portfolio companies of Bridges Fund Management:

- **Fitness Hut** is a high-quality, low-cost Portuguese gym group.
- **Unforgettable** is an e-commerce platform supporting those living with dementia and their carers.
- **The Vet** is a veterinary services provider offering affordable, quality and convenient care to domestic pets across the UK.
- **The Hub Pharmacy** is a full-service community pharmacy group in the UK.

FIGURE 4 Sector and country breakdown of respondents



The survey questions

What data can be collected efficiently across the five dimensions of impact directly from people experiencing change?

To answer this question, we developed two surveys: one for customers and one for employees.

Customer survey

What impact are customers experiencing?

The aim of the survey is to better understand (directly from customers):

- What outcome(s) the product or service is contributing to
- Who the customers are, in terms of demographics and how underserved they are in relation to the outcome(s)
- The degree of change towards the outcome (How Much — depth)
- The duration of the period for which the stakeholder experienced the outcome (How Much – duration), and how quickly it occurs (How Much — rate)
- The Contribution of the product/service to the outcome, relative to what would likely happen anyway
- The impact Risk that the positive impact sought does not occur, and/or that negaitve impact also occurs as a result of the product/service

The customer survey below was used for all of the companies. The first four questions needed re-phrasing for the different survey techniques, as shown below. From question five onwards the phrasing was the same.

#	Dimension	Online survey	Question type	Phone survey	Question type
1	What	When you bought this [product/service], what improvement in your life were you looking for?	Open-ended	Has your quality of life changed because of the [Org] [product/service]?	Close-ended (likert scale[4])
RO1[5]		Write in			
2	How much - depth	To what degree have you experienced this change?	Close-ended (likert scale)	What in your life has changed?	Open-ended
RO1		Improved a lot			
RO2		Improved a little			
RO3		No change			
RO4		Got worse			
RO5		Got a lot worse			
RO6		Don't know			
3	What	How important is this change to you?	Close-ended	How important are each of these changes to you?	Close-ended
RO1		Very Important			
RO2		Important			
RO3		Not very Important			
RO4		Not important at all			
RO5		Don't know			
4	What	Did anything else in your life improve that you think is important?	Open-ended		
RO1		Write in			

[4] The Likert scale asks respondents to select a rating on a scale that ranges from one extreme to another, such as "strongly agree" to "strongly disagree."

[5] 'RO' = Response Option

#	Dimension	Question - methodology agnostic	Type of question
5	What / Risk	Did anything bad happen because of [Org] that is important? If so what?	Open-ended
RO1		Write in	
6	What	Is the change you are experiencing sufficient to meet your needs?	Close-ended (multiple choice)
RO1		Yes, definitely	
RO2		Maybe	
RO3		No, not required	
RO4		No, not satisfied	
RO5		Don't know	
7	How Much - rate	How soon after using [product/service] did these changes happen?	Close-ended (multiple choice)
RO1		Immediately	
RO2		Within days	
RO3		Within weeks	
RO4		Within months	
RO5		Within a year	
RO6		Longer than a year	
RO7		Not sure	
8	How Much - duration	Have these changes been long-lasting?	Close-ended (multiple choice)
RO1		Yes, the changes have been long-lasting	
RO2		No, they stopped after a while	
RO3		No, they stopped too soon	
RO4		It's too soon to know	
RO5		Not sure	
9	Enterprise Contribution	Is there a good alternative to the [Org product/ service] that will deliver the life improvements you want?	Close-ended (multiple choice)
RO1		Yes, but [product/service] is better	
RO2		Yes, but the alternative is exactly the same as [product/service]	
RO3		No	
RO4		Don't know	
10	Enterprise Contribution	Apart from the [product/service] did anything else contribute to the changes you mentioned?	Open-ended
RO1		Write in	
11	Drop-off Risk	Do you intend to buy from [Org] again?	Close-ended (multiple choice)
RO1		Yes, definitely	
RO2		Maybe	
RO3		No, not required	
RO4		No, not satisfied	
RO5		Don't know	

#	Dimension	Question	Type of question
12	Stakeholder participation Risk	Do you expect that [Org] will use the feedback from this survey to improve its work?	Close-ended (multiple choice)
RO1		Yes, definitely, [Org] has been responsive to me before	
RO2		Yes, even though I have not seen [Org] make improvements before	
RO3		Maybe, even though I have no real basis to think it will	
RO4		No, because [Org] has a reputation for not listening	
RO5		Don't know	
13	Who	On a scale of 1 to 5 (where 1 is not at all, and 5 is completely) to what extent do you have the resources available to meet your [insert sector] needs?	Close-ended (multiple choice)
RO1		Not at all	
RO2		Very little	
RO3		Somewhat	
RO4		Mostly	
RO5		completely	
14	Who	On a scale of 1 to 5 (where 1 is much worse, and 5 is much better) how would you say you were doing compared to people around you?	Close-ended (multiple choice)
RO1		1 – much worse	
RO2		2 – slightly worse	
RO3		3 – the same	
RO4		4 – a little better	
RO5		5 – much better	
15	Execution Risk	Have you ever experienced any challenges with [Org product/ service]?	Close ended
RO1		Yes	
RO2		No	
16	Execution Risk	If yes to 15, do you agree with this statement: "[Org] made it easy to resolve my challenge"?	Close ended
RO1		Strongly disagree	
RO2		Disagree	
RO3		Neither disagree or agree	
RO4		Agree	
RO5		Strongly agree	
17	Who	In addition to the above, we asked some specific demographic questions relevant to the company's product/service to help them better understand their customers. For more information on this see the Who dimension section on p.20.	Various

Examples and guidance for using the customer survey questions

Detailed guidance on using the customer survey questions:

Lessons from using the customer surveys

This section maps the survey questions to the data categories for each dimension of impact that were developed through the consensus-building efforts of the IMP, conducting testing and feedback with groups of enterprises and investors. The data categories provide practical guidance on the type of data needed to assess any social or environmental impact.

Whether an enterprise is starting impact measurement from scratch, or whether some data has already been collected, the data categories are a helpful checklist to ensure none of the necessary data-points are missing.

How to navigate this section

This section is organised by the five dimensions of impact.

For each dimension, we present:

- the survey questions piloted for each dimension,
- example responses to the questions from the pilot,
- guidance on how each question can be analysed effectively,
- context on why the resulting data is useful for impact management decision-making, and
- a summary of lessons learned in the pilot about the effectiveness of each question and tips for implementation.

The What dimension

The What impact dimension considers what outcomes the enterprise contributes to — and how important these are to stakeholders.

For the What dimension there are three data categories that enterprises can use to collect, assess and report data on (see Table 3).

TABLE 3 The What dimension of impact

Impact data category	Description of data needed	Question(s) tested
Outcome in period	The level of outcome experienced by the stakeholder when engaging with the enterprise. The outcome can be positive or negative, intended or unintended.	Q1. When you bought this [product/service], what improvement in your life were you looking for? Q.5. Did anything bad happen because of [Org] that is important? If so what? [This question helps uncover unintended negative outcomes. For analysis of this question, see Risk on p.35]
Importance of outcome to stakeholders	Stakeholders' view of whether the outcome they experience is important.	Q3. How important is this change to you? Q4. Did anything else in your life improve that you think is important?
Outcome threshold	The level of outcome that the stakeholder considers to be positive/good enough. The outcome threshold can be a nationally or internationally-agreed standard.	Q6. Is the change you are experiencing sufficient to meet your needs? [The outcome threshold is usually a local or national standard (e.g. a local minimum wage or biometric threshold for healthy). However, it is also useful to ask whether this standard is sufficient for the person affected. We asked this question in conjunction with asking about the degree of change (How Much - Depth), so for analysis of this question, see p.24]

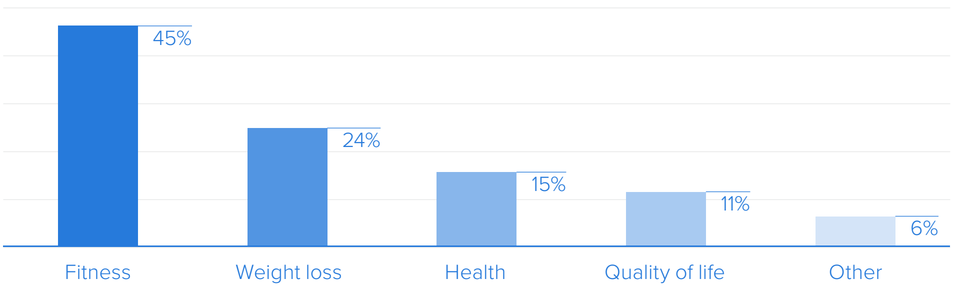
Q1.

When you bought this [product/service], what improvement in your life were you looking for?

By asking respondents what in their life has changed as a result of their engagement, the enterprise can understand which changes are most important from the customers' perspective. This has several purposes; for example it may help an organisation to test the accuracy of their theory of change, or better understand their alignment toward the SDGs. As the question had open-text responses, Keystone worked through the results to group the responses by outcome theme (this is known as 'coding'). To understand which outcomes were most important to manage, these results were considered alongside the results of question 3 (overleaf).

FIGURE 5

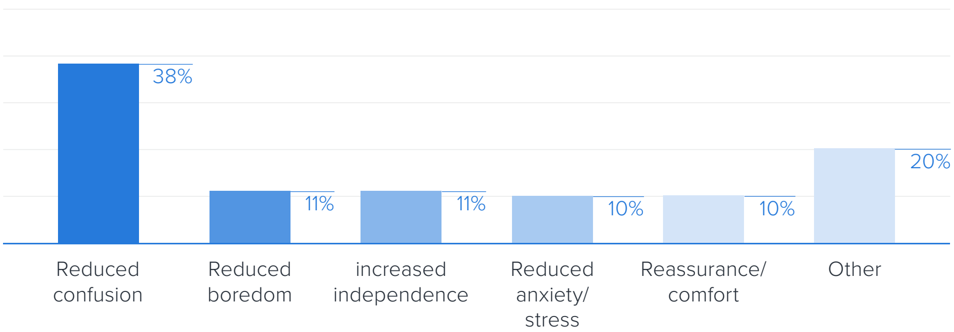
When you first visited [Org], what improvements in your life were you looking for?



For example, the findings for this question from a healthcare enterprise above illustrates that improved fitness was the most common response, with other members seeking weight loss, health or more general quality of life improvements.

FIGURE 6

When you brought a product from [Org], what improvement in your life (or the life of another person) were you looking for?



For example, a healthcare enterprise had found it particularly difficult to define the outcomes sought by their customers. This question was valuable to help them understand how customers perceive what improvements their products cause. Keystone broadly clustered the responses around whether the product had a positive impact on the customer by either reducing confusion (for example help people tell the time, or use a radio), reducing boredom, reducing anxiety/stress, increasing independence, or enabling increased reassurance/comfort for carers themselves.[6]

[6] These are categories that Keystone deemed useful for the purpose of analysis. It may be that with a greater understanding of the business and its customers over time, a healthcare enterprise would find it more useful to code the open text questions into different themes.

Q3.

How important is this change to you?

In addition to information on what had changed, we also sought to understand how important these changes were. This helps the enterprise learn what matters most to their stakeholders so they can focus on improving those outcomes.

For example, an asset-backed financial inclusion company found that the ability to manage household expenses was a more frequently cited important outcome than improved earning potential. A traditional theory of change may have pegged the latter higher.

TABLE 4

	Top improvements in quality of life, by sector		Changes in customers' words	% feel changes "very important"
Education	65%	Improved communication skills	"I had stage fright before...but now my communication skills have improved"	26%
	45%	Improved employment potential	"I have more opportunities. I can work in more fields"	84%
Financial Inclusion	78%	Greater economic resilience	"I do pay school fees on time so my son is never sent home"	86%
	60%	Improved earning potential	"I started my own business and so I do not have to depend on my husband"	89%
Energy	46%	Timely and reliable power supply	"We are getting light energy on time"	13%
	29%	Increase in profit, business hours	"Earlier, we would be forced to shut the shop early, but now we can stay there till late"	14%

Q4.

Did anything else in your life improve that you think is important?

This question seeks to examine to what extent, through their engagement with the enterprise, customers experienced other unexpected positive outcomes that affect their wellbeing, and may inadvertently affect their consumer experience.

For example, customers of a healthcare enterprise attributed a variety of benefits to their membership:

- 23% of respondents mentioned included improvements around better mental health, such as lower stress levels, improved self-esteem, and improved mood.
- 28% of respondents mentioned physical benefits such as increased mobility, feeling more energetic, sleeping better, and increased strength.
- 8% of respondents mentioned improved diet. One person quit smoking.
- 5% of respondents referred to the social benefits, such as making new friends.

The survey also asked about unexpected negative outcomes (see p. 37).

How to analyse this data:

1. Code[7] open text responses by grouping them by whether the answer falls into the same or similar themes.
2. Compare popular themes across different questions to see what customers mention the most.

- After this has been done once, these same themes can inform the multiple-choice options that the company uses for for follow-up surveys. Do keep an “other, please explain” category in case themes change over time.
- Other survey questions will need to be segmented or disaggregated by these same themes to carry out analysis (e.g. the How Much questions on p.24).

Why these questions are valuable for impact management:

- These questions provide enterprises with an insight into the different outcomes delivered by the enterprise, which have been identified by their customers as important to them.
- This data enables enterprises to prioritise and better manage these outcomes - and ultimately improve the impact they are having on customers and other stakeholders.

What we learned

In general, the questions worked very well in terms of getting high quality data on the What dimension. For some sectors and types of product/service, it was natural for the respondent to identify the changes in terms of outcomes sought. For example, as seen in Figure 5, outcome-focused responses came naturally for users of a gym or attendees of a training course.

Where customers did not phrase their answers in terms of outcomes, this affected the usefulness of many other questions. There were two cases where customers did not phrase their responses in terms of the outcome sought:

- Customers of a pharmacy were less likely to respond in terms of the outcome they sought and instead listed the activity the pharmacy enabled (e.g. prescription collection).
- Customers of a vet similarly responded with the value proposition of the product (e.g. cost or convenience) rather than the outcome they hoped for from the vet.

In cases like these where the survey did not surface clearly-defined outcomes, the answers to the subsequent questions were less valuable, as all the questions relate to the outcome articulated (e.g. the depth or duration of outcome experienced).

Our recommendation for cases like these would be to include multiple-choice response fields alongside free text fields, in order to encourage responses to be outcome-focused. In the case of a pharmacy, for example, a fixed multiple-choice list of types of ailments could be provided, alongside questions on the value proposition of this pharmacy versus another.

[7] For a simple guide to coding see http://onlineqda.hud.ac.uk/Intro_QDA/how_what_to_code.php

The Who dimension

The Who impact dimension considers which stakeholders experience the social and environmental outcomes.

The data categories under the Who dimension help enterprises identify the characteristics of the stakeholders they affect — and understand how underserved they are in relation to the social or environmental outcomes sought.

TABLE 5 The Who dimension of impact

Impact data category	Description of data needed	Question(s) tested
Stakeholder	The group experiencing the outcome	*No survey question necessary – enterprise can easily identify 'customer' vs 'employee' stakeholder groups.*
Geographical boundary	The location in which the stakeholder experiences the social and/or environmental outcome	Q17+. Specific demographic question on customer location. [This data can be used to segment results obtained across the other dimensions and uncover trends.] E.g. Which store location do you visit?
Stakeholder characteristics	Socio-demographic (e.g. age, gender, ethnicity) and behavioural (e.g. motivations and beliefs) characteristics of the stakeholder relevant for segmentation	Q17+. Specific demographic questions on customer demographics. [This data can be used to segment results obtained across the other dimensions and uncover trends.] E.g. What is your age? What is your gender? What department do you work in? Poverty/ need level in relation to outcome
Baseline outcome level	The level of outcome experienced by the stakeholder prior to engaging with the enterprise	Q13. On a scale of 1 to 5 (where 1 is not at all, and 5 is completely) to what extent do you have the resources available to meet your [insert sector] needs? Q14. On a scale of 1 to 5 (where 1 is much worse, and 5 is much better) how would you say you were doing compared to people around you?

Q17+

Specific demographic questions were used to collect data on geographical location and stakeholder characteristics:

Demographic questions provide enterprises with a better understanding of who their customers are. Then, by segmenting the outcome data (from questions 1-5) by each demographic group, the company can learn whether the customer experience differs for different market segments (e.g. young vs old, female vs male etc.).

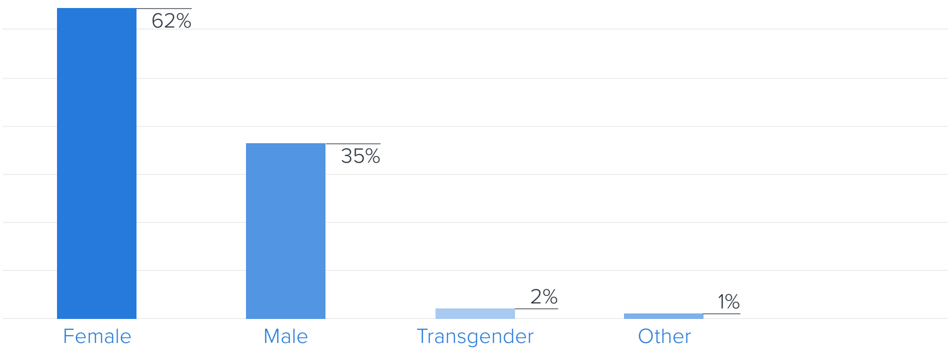
For impact-focused businesses, this data can help ensure the most underserved customers are reached. This is effective when customer demographics or location are a good indicator of their level of need of the outcome (e.g. poverty level is often a good indicator of a high level of need of some outcomes - but not all).

Given that the relevant demographic information differs by enterprise, when selecting these questions it is often helpful to consider:

- What data do you already hold on your customers that might be relevant for usefully disaggregating responses? For example – do you already have data on where they live, how old they are, or how often they use your product/service?
- What do you want to know about your customers? For example - do you want to know whether men and women experience your services differently?

FIGURE 7

What is your gender?



For example, The Hub pharmacy wanted to know the gender of their clients in order to learn whether there were patterns of engagement with the pharmacy based on gender, and whether this should influence marketing.

When analysing the survey results, the team therefore segmented the responses from questions 1-5 by these demographic results to analyse trends.

Q13.

On a scale of 1 to 5 (where 1 is not at all, and 5 is completely) to what extent do you have the resources available to meet your [insert sector] needs?

Q14.

On a scale of 1 to 5 (where 1 is much worse, and 5 is much better) how would you say you were doing compared to people around you?

These questions seek to understand the baseline data about level of need. This data is particularly important in situations where demographic data doesn't necessarily help with understanding the customer's level of need in relation to the outcome of a particular intervention or business model (e.g. mental health).

The timing of these questions matters. Ideally, they would be asked before the intervention of a product or service to establish a baseline level of need, and then asked again after customers have had enough time to experience the outcome of the product or service, to understand improvements. Without being able to compare the baseline and endline, it is difficult to determine the enterprises' effect on the level of need of a customer.

FIGURE 8

Q.13 On a scale of 1 to 5, how would you say you were doing compared to the people around you?

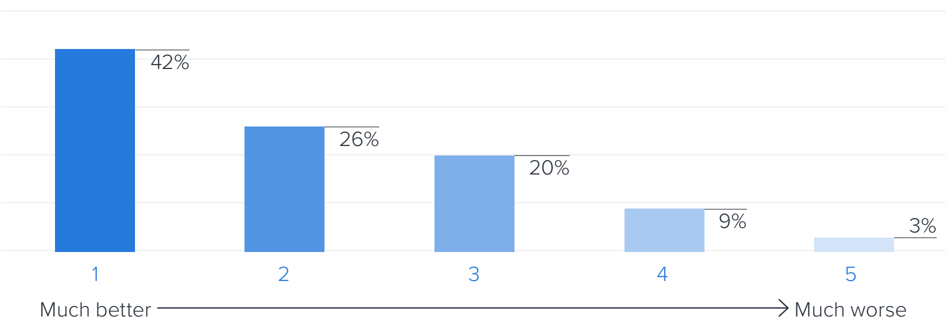
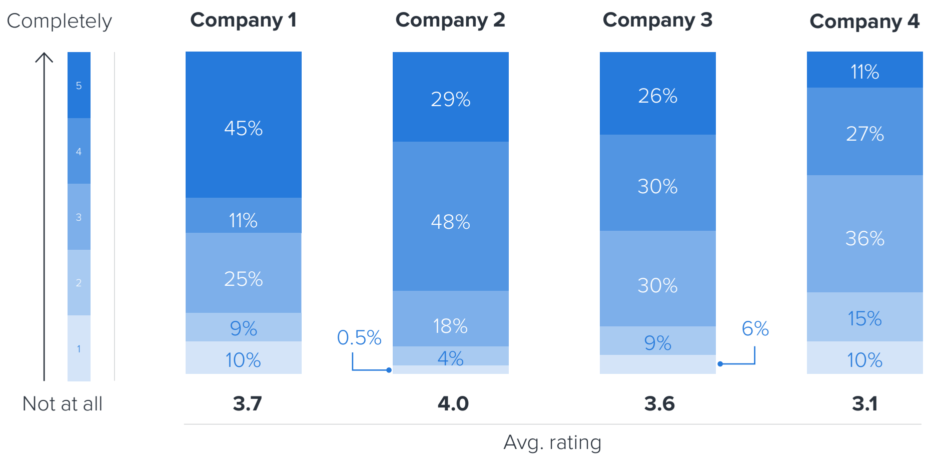


FIGURE 9

Q.14 On a scale of 1 to 5, to what extent do you have the resources to meet your [insert sector] needs?



How to analyse this data

- Segment the results from the other questions by relevant demographic groups.
- Analyse results to identify trends or patterns.

Why these questions are valuable for impact management:

- Data on the Who dimension is a vital part of impact management, as it not only helps an enterprise understand more about the demographics and location of the stakeholders it affects, but also enables impact-driven enterprises to ensure they're delivering outcomes for the most underserved.
- This data is particularly important for enterprises that are seeking to contribute to the SDGs, or that are trying to 'Contribute to Solutions' more broadly[8].

What we learned

The questions developed in this survey worked well when asking for specific, easily identifiable demographic or geographic characteristics of customers (e.g. Q17).

It was more challenging to collect accurate demographic data where the product or service was not used by the respondent. For example, some people would go to a pharmacy to pick up a prescription for a family member or a friend. Collecting demographic or identifying information about the characteristics of that person is not as relevant as collecting that data from the person actually needing the prescription.

Where demographic or geographic characteristics are a good indicator of whether the customer is likely to be underserved in relation to the outcome, these questions are effective. And there are several other effective established surveys that can be used in addition to these questions. For instance, the Poverty Probability Index is a helpful, simple way to measure the poverty levels of stakeholders.

Where it is not relevant to use demographic or geographic data to assess how underserved a population is in relation to the outcome, an additional question is needed to understand the level of need at the baseline. The questions we tested to understand the extent to which a customer is underserved in relation to the outcome sought (Q13 and 14) performed less well in the pilot and may need to be re-phrased and asked only at baseline. Asking someone whether they have the resources to meet their needs after engagement with the enterprise doesn't uncover prior need and could be affected by bias.

[8] 'Contribute to Solutions' is one of three enterprise impact categories used to segment the investment market based on observable impact performance across the five dimensions. A benefit of classification is that could enable more appropriate matching of investor intentions with available investment product (or enterprises). Read more here: https://impactmanagementproject.com/impact-management/how-investors-manage-impact/

The How Much dimension

The How Much dimension considers the significance of the outcome — across scale, depth, and duration.

The data categories under the How Much dimension help enterprises measure how much of the outcome has occurred.

TABLE 6 The How Much dimension of impact

Impact data category	Description of data needed	Question(s) tested
Scale	The number of individuals experiencing the outcome. When the planet is the stakeholder, this category is not relevant.	*No feedback question necessary – enterprise can collect scale data through other means more effectively.*
Depth	The degree of change experienced by the stakeholder	Q2. To what degree have you experienced this change? Q6. Is the change you are experiencing sufficient to meet your needs?
Duration	The time period for which the stakeholder experiences the outcome	Q8. Have these changes been long-lasting?
Rate	The time taken from the point of engagement with the enterprise to the outcome being experienced	Q7. How soon after using [product/service] did these improvements happen? [IMP has had feedback that this sub-dimension is not relevant to many enterprise models or impacts, and so is optional]

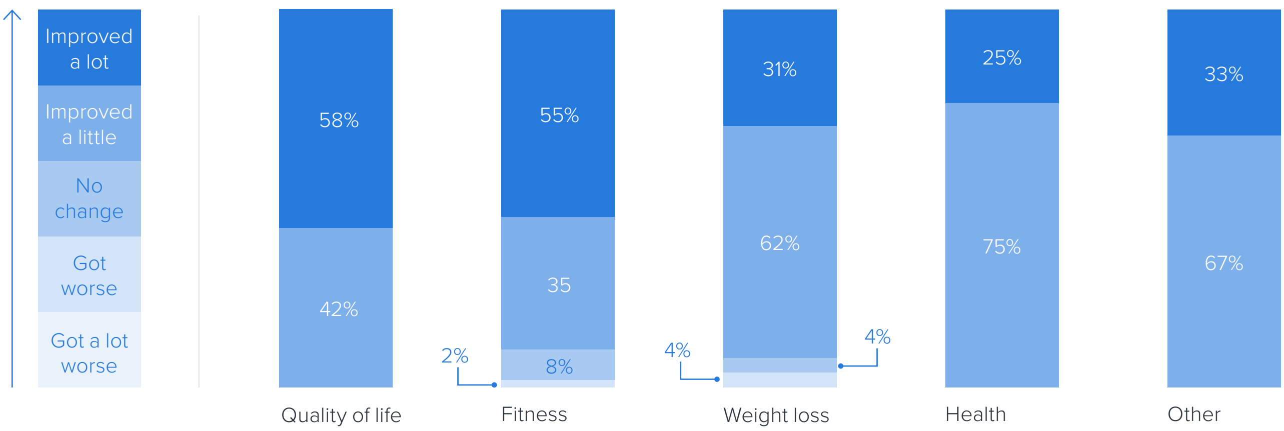
Degree of change / depth

Q2.

To what degree have you experienced this change?

Question 2 asks respondents to determine the degree to which they experienced the outcome(s) articulated in questions 1-5.

FIGURE 10



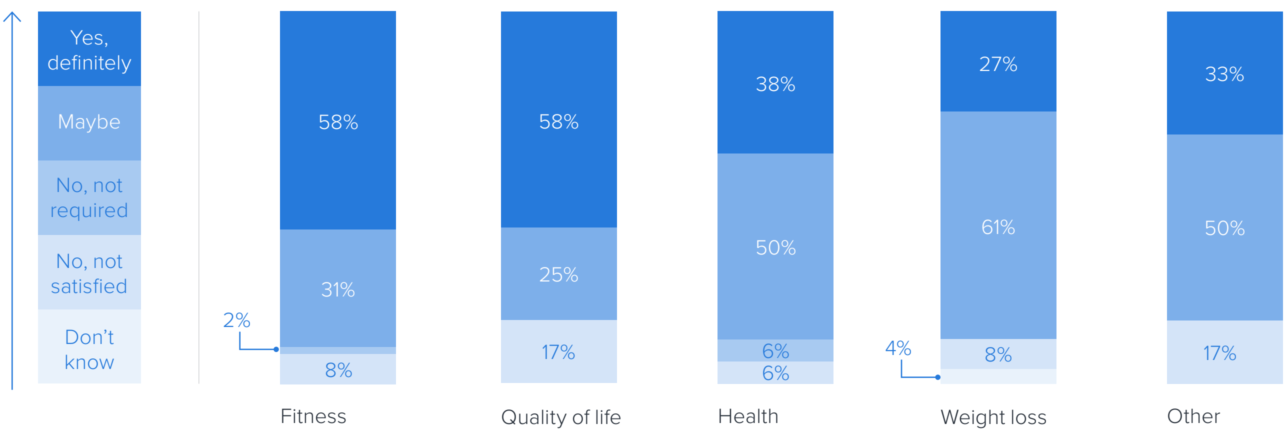
Q6.

Is the change you are experiencing sufficient to meet your needs?

Question 6 seeks to determine whether the change reported is sufficient to meet the respondent's needs.

The degree of change for someone seeking weight loss or an improved quality of life might be dependent on a number of factors external to the product or service. Although this question cannot determine causality, the customer perception is what will matter to the enterprise - which can use this data to manage the customer experience.

FIGURE 11



For example, the enterprise can use this feedback to engage in a dialogue with respondents to gain a better idea of customer expectations, and what the business could be doing to improve the results against these expectations.

How to analyse this data:

1. Segment responses to both questions by the outcome reported by respondents in questions 1-5
2. Consider which outcomes are perceived by customers to be driven most deeply by the product/service as a basis for management decision-making to improve the product/service

Why these questions are valuable for impact management:

- Depth data helps an enterprise determine how effective their product/service is by considering the degree of improvement towards the outcome that has occurred since the customer's initial engagement with the product/service (i.e. the baseline)
- This question also helps to uncover not just whether there has been a change in the degree of outcome experienced, but whether the customer attributes this change to the enterprise.

What we learned

This question performs well at giving a first impression of depth. Enterprises reported that they found it both intuitive and informative.

This question is particularly important when non-self-reported objective measurement of the degree of change experienced is challenging. For example, this self-reported depth data was incredibly valuable for the gym and dementia enterprises, both of which otherwise struggle to track increased outcomes through non-self-reported data.

However, detailed assessments of depth of impact will usually require further analysis, using sector-specific questions. For an example of this in practice, see the Acumen Energy Report (p.28) which features a numbers of standardised, outcome-level measures used to calculate depth of impact across an off-grid energy portfolio.

Duration

Q8.

Have these changes been long-lasting?

The majority of the customers 60 Decibels spoke to said the changes in their lives had been long-lasting. While it may be the case that all of the changes that occurred in their lives since engaging with the company had endured, with hindsight we consider that this question may be slightly leading. It could be more objective to ask more broadly about the duration of the changes, i.e. "How long have the changes due to the company endured?"

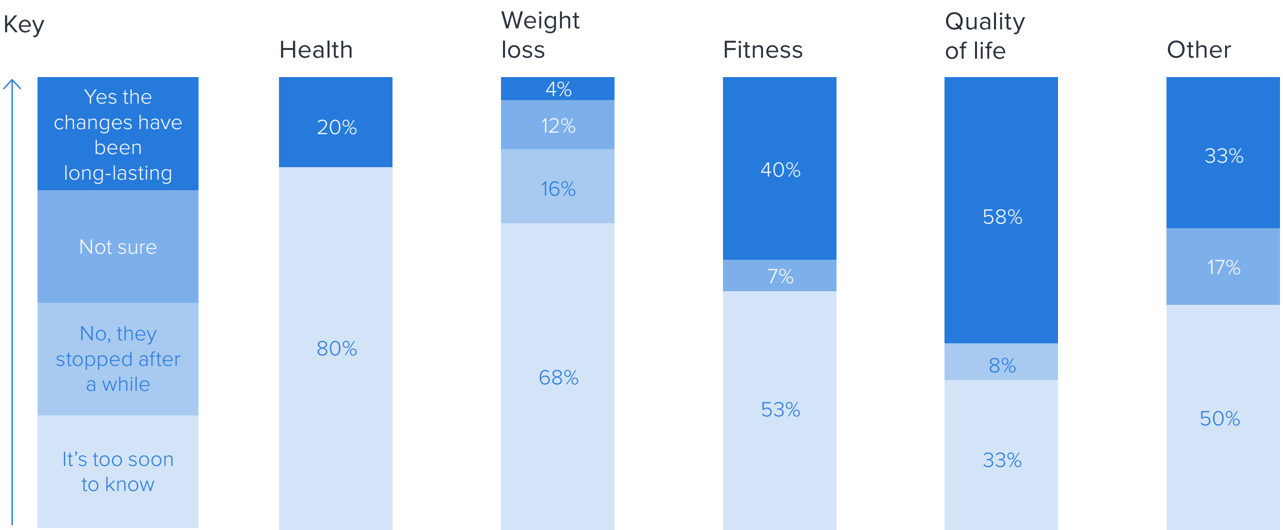
FIGURE 12

Have the changes you've experienced been long-lasting?

However, as shown by Figure 13 below, most respondents of a healthcare enterprise felt that it was too soon to tell whether the improvements they are seeking and experiencing are or have been long-lasting. This is partially due to the diverse improvements being sought but also the values of the individual. For instance, someone who is highly dedicated to losing weight and has managed to do so by using the product/service will only be able to tell whether this weight loss is sustainable over the coming years. Someone who is attempting to be healthier may experience this improvement right away, but the extent to which this lasts is very much dependent on the individual and their commitment to this improvement.

FIGURE 13

Have these changes been long-lasting?



How to analyse this data:

- Responses to this question also need to be segmented by the outcome articulated in questions 1-5.
- Responses can also be cut by demographic information (see Who on p. 20) to uncover trends
- It is challenging to remove bias caused by the time point at which this question is asked in the customer journey, and therefore knowing the point in time that the question is asked is very important (beginning, middle or end, for example). For each type of enterprise there will likely be an optimum point in time to ask this question to ensure outcomes are proving long-lasting.
- The enterprise needs to know what the expected duration of change is to interpret these results fully and improve their product/service's impact as a result.

Why these questions are valuable for impact management:

- How long an outcome lasts for is a particularly difficult dimension of impact to measure with non-self-reported data for many enterprises and products/services, so surveying can be a useful method for gathering data on this dimension in particular.
- This question has the added benefit of helping an enterprise assess drop-off and endurance risk (read more on Risk on p.35)

What we learned

Take into account the expected duration. We concluded that it was important to tailor this question to the specific expected duration of each business model. For example, if someone is going to the vet because their dog has a broken leg, the outcome of having the vet care and helping to nurse it back to health, generally has a given timeframe of several weeks/months. However, if someone with dementia is looking for support on something like decreasing pain or improving agility, this is highly context-specific and sometimes there is no timeline to such an outcome. Someone's perception of an outcome of 'good health' therefore changes based on context and their expectations from a particular product, service or intervention.

Don't ask leading questions. It is likely that the phrasing of the response options made this question too leading, as the majority of respondents replied, "Yes, the changes have been long-lasting." It would be advisable to slightly alter the question to a more open-ended framing to prevent bias (e.g. "How long have the changes endured?).

Factor in whether the resulting duration meets the customer's level of need, where possible. In a similar way to the questions on degree/depth, this question becomes most actionable for an enterprise if the response includes feedback on whether the duration of the outcome was as expected and/or sufficient to meet the client's needs. In future, we intend to test a question that asks how long the client would like the outcome to last (asked at baseline) and then test a follow-up question asked at regular intervals relevant to the enterprise model where the client is asked whether the duration of the outcome has met their needs.

The longer the duration, the harder the attribution. Attributing change to the enterprise becomes more challenging as time goes on and other external factors intervene. As a result, most enterprises will not be able to reliably attribute the enduring outcome to the effectiveness of their product/service alone.

Rate

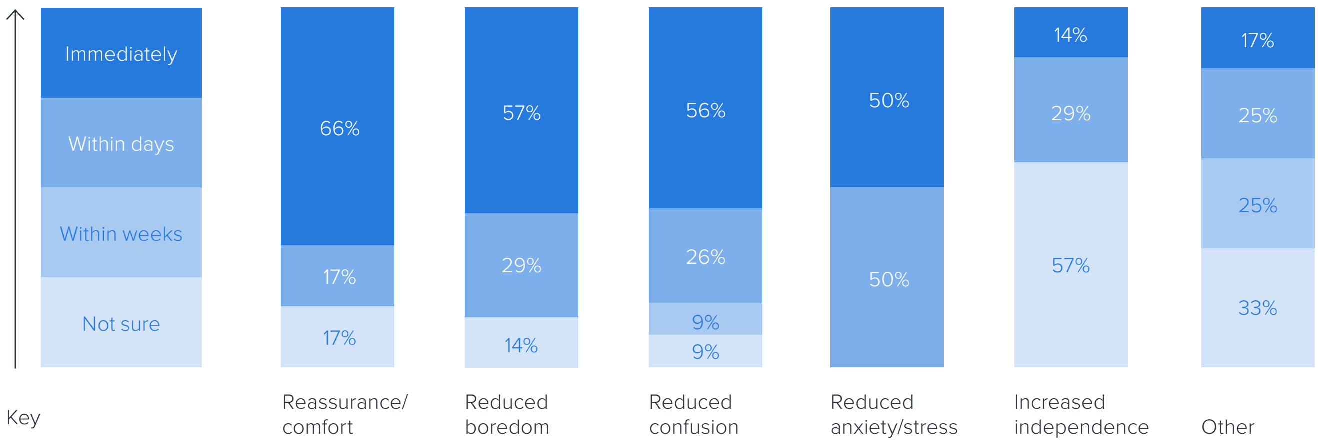
Q7.

How soon after using [product/service] did these improvements happen?

Similar to the question addressing how long-lasting a change has been, asking respondents to evaluate how quickly a change has occurred, is dependent on the type of change being sought and the commitment of the individual to achieve this change. Again, the business will be better able to use this data to improve performance if they have a sense of the expectations for each outcome.

FIGURE 14

How soon after using [product/service] did the changes happen?



For example, given that those experiencing 'reassurance', 'reduced confusion' and 'reduced boredom' tended to experience the outcome immediately, those products and services look to be most effective at enabling quick improvements.

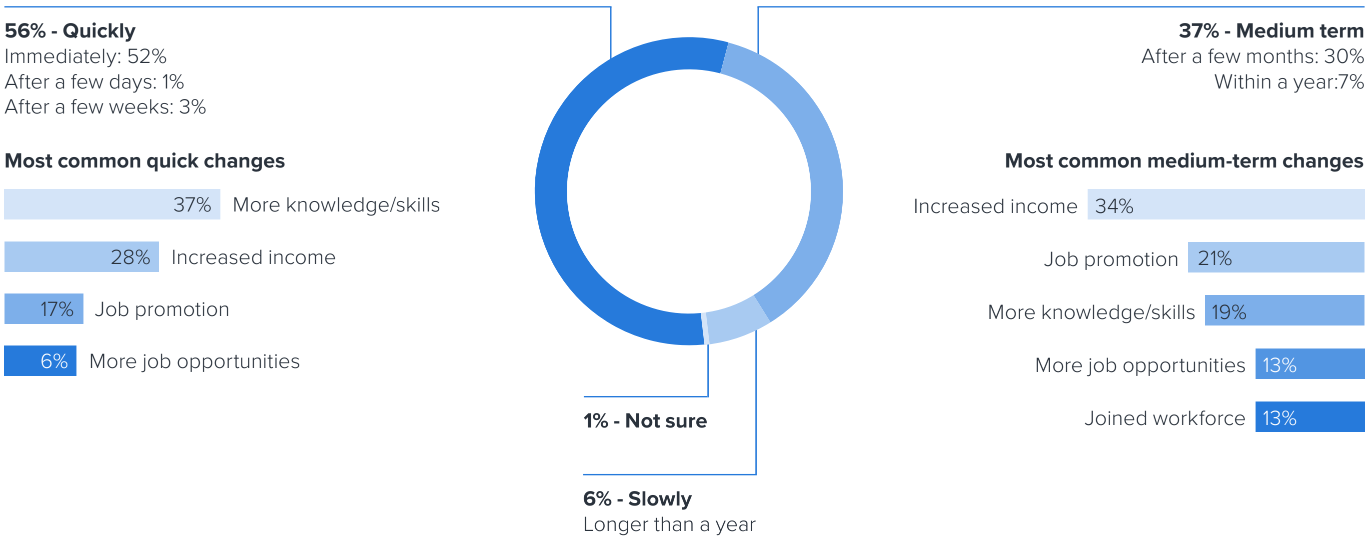
In another example, from a vocational education company:

- Almost 60% of clients felt that changes started within weeks of graduating, while almost everyone else said it set in within months.
- The results illustrate how the impact unfolds over time. The top immediate change is improvements in skills / knowledge, while the top medium-term change (i.e. within a few months to a year) is increased income.

These examples illustrate how different products and services have different gestation period in terms of the time it takes impact to be felt by the beneficiary. So understanding customer expectations, and setting goals accordingly, is important for this question.

FIGURE 15

How soon after graduating from [Org] did the changes happen to you?



How to analyse this data:

- Responses to this question also need to be segmented by the outcome articulated in questions 1-5.
- Responses can also be segmented by demographic information (see WHO on p. 20) to uncover trends
- However, as with duration, what constitutes 'good performance' for the rate dimension is highly subjective and often dependent on the type of improvement sought.
- In addition, other external factors often come into play that affect the enterprise's ability to manage performance for this dimension. For example, the rate at which a gym-goer achieves their desired outcome of good health is determined partially by the effectiveness of the gym and partially by external factors, such as each individual's dedication, genetic disposition, and feasible time investment.

Why these questions are valuable for impact management:

- Unlike the other How Much sub-dimensions, the rate question may only be relevant to some effects or business models. For example, a business might be particularly interested in this dimension if speed is their unique selling point (e.g. the customer will receive quicker treatment or learn a skill faster) or if the outcome is dependent on the speed of delivery (e.g. life-saving medical care or disaster relief).

What we learned

It is not always relevant to collect data on rate for every type of impact. For example, this question was not effective when seeking to understand the impact of an online information service - the rate of access was immediate for all users, and the speed of information absorption was not a meaningful metric. Rate is also less actionable data for a company where the customer's own choices and behaviour play a bigger role (e.g. improved health through gym membership).

The timescale used in the survey responses for duration and rate needs to be based on the customers' expectations for a particular outcome. This helps ensure that the responses to the questions are actionable for the company. We therefore concluded that these responses could not be standardised across outcomes, as what constitutes a realistic and desired rate and duration of change will vary by outcome (e.g. learning a new skill versus receiving urgent medical care).

The Contribution Dimension

The Contribution impact dimension recognises that impact occurs in a dynamic social system, with various stakeholders playing different roles.

To understand their own contribution to a social or environmental outcome, enterprises need to consider what would have happened in absence of their activities. The data categories under the Contribution dimension help enterprises and investors assess an enterprise's contribution to the social (environmental) outcomes that people (planet) experience, relative to what the market or social system would have done anyway.

TABLE 7 The Contribution dimension of impact

Impact data category	Description of data needed	Question(s) tested
Depth counterfactual	An enterprise's contribution to the depth of an outcome, by factoring in the estimated degree of change that would have otherwise occurred for the stakeholder.	[These questions consider the depth and duration counterfactual together] Q9. Is there a good alternative to the [Org product/service] that will deliver the life improvements you want? Q10. Apart from the [product/service] did anything else contribute to the changes you mentioned?
Duration counterfactual	An enterprise's contribution to the duration of an outcome, by factoring in the estimated time period that the outcome would have otherwise endured for the stakeholder.	

The survey tested two ways of assessing enterprise contribution:

Q9.

Is there a good alternative to the [Org product/service] that will deliver the life improvements you want?

FIGURE 16

Could you easily find a good alternative to [company]?

This question helped us figure out whether the company was addressing a market gap (e.g. the 52% for this company). This question helps the enterprise determine whether the likely impact might have happened without them.

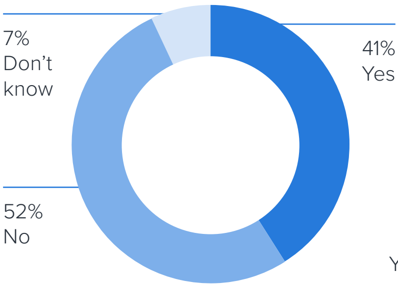
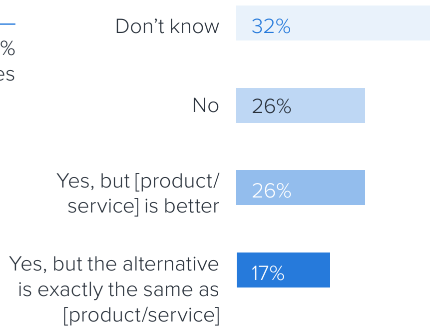


FIGURE 17

Is there a good alternative to [company] that will deliver the life improvements you want?



With the online survey, the responses to the question were more detailed and sought to test more specifically how the impact compares with what a competitor business might deliver for the same target market.

These results were inconclusive.

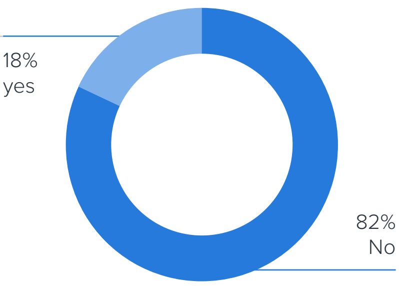
Q10.

Apart from the [product/service], did anything else contribute to the changes you mentioned?

This question seeks to determine whether the same impact is being driven by another source, or whether it is the business alone. This helps the enterprise determine whether the same impact would likely not happen without them.

Over 80% of the customers 60 Decibels spoke to attributed the change that they had experienced to the company. This gives a subjective indication that the company is the primary contributor to the changes observed. However, this data has to be treated with caution. For instance, customers may fail to distinguish the relative impact that several products or services may have to the same outcome. In other words, it may difficult for them to distil which one actually caused the change. For example, a student may complete multiple types of vocational training, or a borrower may take loans from various providers.

FIGURE 18



Judgement will need to be used as to the complexity of causal relationships and likelihood that an individual can intelligently attribute that impact. It will often be the case that to understand contribution, a significantly more rigorous assessment of causal impact will be needed. Feedback, therefore, should be used as a guide, but not necessarily a definitive account of impact.

How to analyse this data:

- Given that these questions invite a customer's perception of whether the enterprise is better than what the market otherwise offers, it is important to substantiate this feedback with market competitor analysis.
- Follow-up questions teasing out why the product/service is better or worse than the next best alternative would help management teams make this data more actionable.

Why these questions are valuable for impact management:

- Impact-driven enterprises with goals to meet an unmet social or environmental need rely on enterprise contribution data to ensure that they are delivering impact that wouldn't otherwise occur. The most accurate way of assessing this dimension is by using a counterfactual. In absence of counterfactual data, this self-reported data can act as an indicative proxy.

What we learned

We concluded that the answer options for Question 9 risked being misleading, and might have steered the respondent in a particular direction that hinders an objective response. In future it would be recommended that the responses to this question include the option: "Yes, and the alternative is better".

There are many other measurement techniques which can be used to measure enterprise contribution and complement the self-reported data collected for this dimension. Other approaches to measuring an enterprise's contribution to the outcome include:

1. **Local benchmark.** To assess what would outcome would likely have happened for a particularly demographic, local government or industry statistics are often a helpful place to start. For example, data might exist on the average wage in that industry, or the prevalence of a disease in a region, or the cost of a specific type of healthcare in a region.
2. **Competitor benchmark.** Where government or industry data is scarce, enterprises can benchmark their impact performance against their closest peer in the market. The enterprise should try to obtain whatever performance data they can about the competitor to evaluate whether their performance is likely better than what is happening elsewhere in the market for that particular demographic group.
3. **Ecosystem analysis.** To obtain a more complete understanding of what else might be contributing to an outcome for a particular demographic, fuller market research can be undertaken to build a picture of the market ecosystem. This assessment looks beyond peer organisations to the contribution of other actors and influences such as: government interventions, external factors (weather, economic conditions), and individuals' unobservable characteristics (self-motivation, cultural practices).
4. **Third-party evidence.** Enterprises can review counterfactual evaluations conducted on similar enterprise models to gather evidence on the likely counterfactual of their enterprise model. However, before extrapolating results from third party evidence, enterprises should ensure that the context of the comparable model is exactly the same as the enterprise's own context, or the evidence will not be reliable.
5. **RCTs and quasi-experimental methods.** Research methods can be used to scientifically measure the difference in outcomes over time using a control group to create a counterfactual. RCTs build this control group using randomisation and other quasi-experimental methods use a range of statistical techniques to build experimental groups.

The Risk dimension

The Risk dimension assesses the likelihood that impact will be different than expected, from the perspective of the people or the planet who experience impact.

There are nine types of risk to guide an enterprise's assessment. Some impact risk factors can be assessed by the availability of the data across other questions. But we included a few questions for five of the risk factors.

The Risk dimension

TABLE 8 The Risk dimension of impact

Type of impact risk	Definition	Question(s) tested
Evidence risk	The probability that insufficient high-quality data exists to know what impact is occurring (or will occur) across the other four dimensions of impact, for all stakeholders.	*No question was tested. This risk factor is best assessed by looking at the availability of good quality data across the other dimensions.*
External risk	The probability that external factors disrupt our ability to deliver the expected impact.	*No question was tested. This risk factor is best assessed by considering factors external to the business model.*
Unexpected impact risk	The probability that significant unexpected positive and negative impact is experienced by people and the planet.	Q5. Did anything bad happen because of [Org]? If so what?
Drop-off risk	The probability that the expected positive impact does not endure and/or that negative impact is no longer mitigated.	*No question was tested. This risk factor is best assessed by considering the availability of good quality data on the duration dimension.*
Efficiency risk	The probability that the expected impact could have been achieved with fewer resources or at a lower cost.	*No question was tested. This risk factor is best assessed by looking at the resources used to deliver impact through this business model compared to alternative strategies.*
Execution risk	The probability that the activities are not delivered as planned and do not result in the desired outcomes.	Q15. Have you ever experienced any challenges with [Org product/service]? Q16. Do you agree with this statement: '[Org] made it easy to resolve my challenge?'
Alignment risk	The probability that impact is not locked into the enterprise model, making mission-drift more likely.	*No question was tested. This risk factor is best assessed by identifying which locks exist in the business model to prevent mission-drift (e.g. mission lock, asset lock, performance lock).*
Endurance risk	The probability that the required activities are not delivered for a long enough period.	Q11. Do you intend to buy from [Org] again?
Stakeholder participation risk	The probability that the expectations and/or experience of stakeholders are misunderstood or not taken into account.	Q12. Do you expect that [Org] will use the feedback from this survey to improve its work?

Unexpected Impact Risk questions

Q5

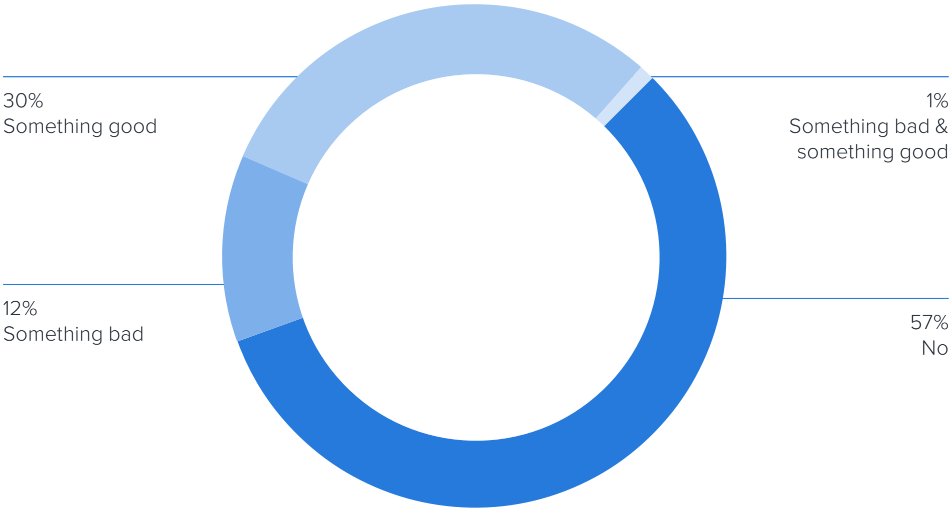
Did anything bad happen because of [Org]? If so what?

This question seeks to assess the likelihood that unexpected negative impact occurs as a result of the enterprise, so that a management team can take action to mitigate this impact. Responses to this question tended to point to very specific problems or disappointments.

As shown in Figure 19 below, the majority of respondents did not report anything bad happening as a result of using the product or service (62%). The remaining respondents listed criticisms with the quality of service: 27% noted a lack of information/confusion, 22% mentioned poor facilities, 18% didn't like the atmosphere, and 16% mentioned overcrowding.

FIGURE 19

Did anything unexpected - good or bad - happen because of [Org]



Execution Risk questions

Q15.

Have you ever experienced any challenges with [Org product/ service]? Describe the challenges you experienced.

Q16.

Do you agree with this statement: '[Org] made it easy to resolve my challenge?'

To understand execution risk, we asked customers on the phone if they had experienced challenges with the company; and if so, what those challenges were and how easily they were resolved. We wanted to understand what aspects of the service were not delivered as planned, which could affect the realisation of the desired outcomes. We also asked whether the company made it easy for customers to resolve their challenges; this helps us to understand how well the company is currently mitigating execution risk.

As shown below, in the case of the microfinance enterprise surveyed, about one-third of customers experienced issues with their loan officer or teacher, including negative customer service experiences, and one-quarter had difficulties with the terms of their group loans. 60% of these customers who experienced challenges did not feel that the company made it easy to handle their issue, indicating that this was an area where the company needed to invest more in order to ensure that its services are delivered as expected.

FIGURE 20

Q15. Please describe the challenges you experienced

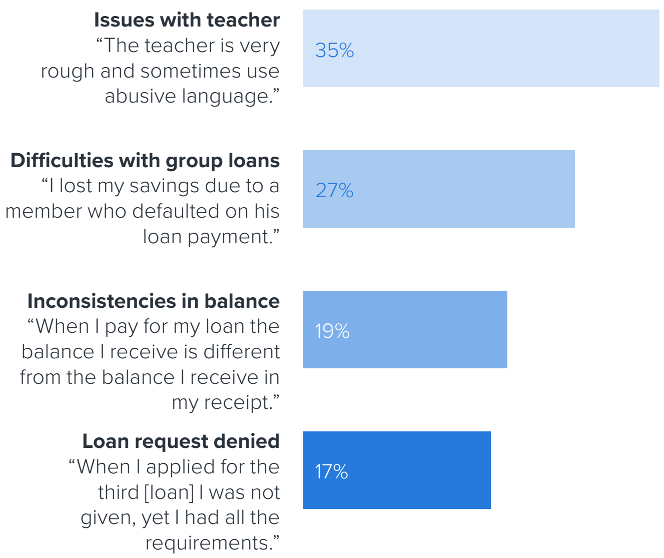
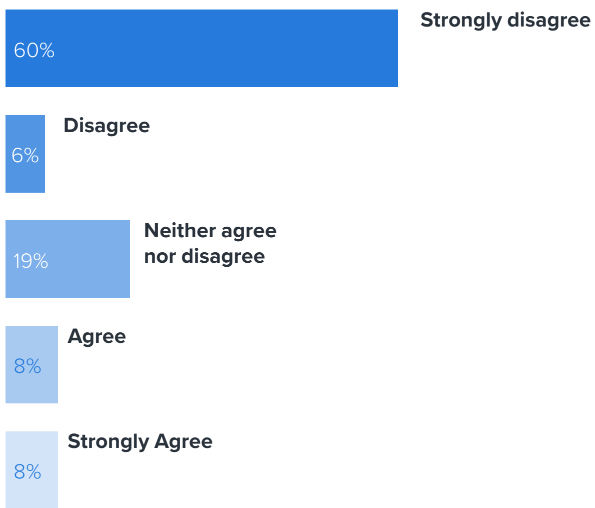


FIGURE 21

Q 16. Do you agree or disagree with this statement: [Org] made it easy for me to handle my issue?



Endurance Risk question

Q11.

Do you intend to buy from [Org] again?

To assess the likelihood that the impact will ensure, we asked customers whether they planned to purchase the company's product or service in the future. This question usually helps to capture the level of customer satisfaction with the enterprise.

However, the context of the specific product or service in question needs to be considered when analysing the results of this question. For example, 21% of graduates of one education company that we spoke to said they did not plan to take another course with the company because it was not required, indicating perhaps that they had the necessary knowledge to do their job, or had the degree they needed to work in their field. By comparison, 92% of loan borrowers said they definitely planned to take another loan from the company in order to continue growing their assets or start new businesses.

Stakeholder Participation Risk question

Q12.

Do you expect that [Org] will use the feedback from this survey to improve its work?

This question tests whether the enterprise respects the views of its customers, as a way of assessing the enterprise's relationship with its stakeholders. A customer's perception of whether a company listens to them and takes their opinion into account is indicative of good stakeholder management, and also suggests the brand has a good reputation among its clients.

The very action of surveying stakeholders, and taking these responses into account when making decisions, also in itself reduces the stakeholder participation risk that the enterprise is taking.

How to analyse this data:

- Given that the availability of impact data and the level of impact risk is closely linked, it is helpful to triangulate responses to these questions with the data available across the other four dimensions.

Why these questions are valuable for impact management:

- Assessing impact risk helps the enterprise know where it needs to focus mitigation, which often involves figuring out what data is needed to improve understanding of the stakeholders' experience. For guidance on assessing each risk factor, see Table 8 below.

TABLE 8

Impact data category	Description of data needed	Illustrative values
Type of risk	The type of risk that may undermine the delivery of the outcome, from the perspective of the people or planet who experience it. Enterprises and investors are exposed to 9 types of risks.	**Execution risk** In order to achieve high absolute learning outcomes rather than high learning progress, schools may select students on the basis of their ability and in the process exclude disadvantaged groups.
Level of risk	The level of risk, factoring in the severity and likelihood of the impact risk.	**High** Market research and expert interviews indicate that this risk is very likely to happen, and could have significant negative consequences in achieving education equality.
Mitigation strategy	Concrete action steps to mitigate the risks identified in the previous two categories. Enterprises and investors should prioritise high- and medium-level risks.	**Mitigants implemented** 1. Schools will be assessed based on progress measures rather than absolute measures. 2. Monetary incentives will be put in place to encourage schools to enrol students from disadvantaged backgrounds. 3. Spot check policy will be implemented to ensure responsible behaviour from the schools.

Guidance for using the employee survey questions

Designing the employee survey

Rather than asking employees free-form questions about what outcomes employment leads to, we first consulted available research about the commonly-agreed drivers of quality employment. The United Nations Economic Commission for Europe (UNECE) identifies seven areas which are listed in Figure 22 below.[9]

The available research suggested that the first five of these areas are best assessed through non-self-reported measurement, examples of evidenced indicators developed by UNECE to measure these five areas are illustrated in the table below.

The Gallup 12-Question Quality Employment Survey is a well-evidenced and widely used series of questions to assess the final two driving factors, which are well-suited to self-reported measures.

FIGURE 22

The seven driving factors of quality employment (adapted from UNECE[11])

Quality employment driving factor	Example UNECE-recommended indicators
1. Safety and ethics of employment a. Safety at work b. Child labour and forced labour c. Fair treatment of employment	*a. Percentage of employed persons who are exposed to physical health risk factors at work* *b. Percentage of children aged 5 to 17 years who are engaged in child labour* *c. Pay gap between subpopulation groups (e.g. gender pay gap)*
2. Income and benefits from employment a. Income from employment b. Non-wage pecuniary benefits	*a. Mean nominal monthly / hourly earnings of employees (local currency)* *b. Percentage of employees entitled to paid annual leave*
3. Balancing work and non-working life a. Working hours b. Working time arrangements b. Work-life balance	*a. Percentage of employed persons working more than one job* *b. Percentage of employed persons who usually work on the weekend* *c. Mean duration of commuting time between work and home (one way)*
4. Security of employment and social protection a. Security of employment b. Social protection	*a. Percentage of employed persons aged 25 years and older with a fixed-term contract* *b. Percentage of employed persons who are active contributors to a pension scheme*
5. Social dialogue	*Days not worked due to strikes and lockouts per 1000 employees (or employed persons)*
6. Skills development and training	Gallup have identified 12 questions that are good indicators for these final two drivers. The survey questions therefore capture the degree to which positive outcomes are occurring in these areas.
7. Workplace relationships and work motivation	

[9] http://www.unece.org/fileadmin/DAM/stats/publications/2015/ECE_CES_40.pdf

The employee survey questions

The aim of the survey is to better understand (directly from employees):

- The degree of change experienced in the two areas of skills and relationships (How Much - depth).
- The Contribution the enterprise makes to decent work relative to what would happen anyway
- When collecting Who data from employees, the same rules apply as for the customer survey (see p. 20).

#	Dimension	Question	Type of question
1	How Much	How far do you agree with these statements?[10]	Close-ended (likert scale)
RO1		I know what is expected of me at work.	
RO2		I have the materials and equipment I need to do my job right.	
RO3		At work, I have the opportunity to do what I do best every day.	
RO4		In the last seven days, I have received recognition or praise for doing good work.	
RO5		My supervisor, or someone at work, seems to care about me as a person.	
RO6		There is someone at work who encourages my development.	
RO7		At work, my opinions seem to count.	
RO8		The mission/purpose of my company makes me feel my job is important.	
RO9		My associates (or fellow employees) are committed to doing quality work.	
RO10		I have a best friend at work.	
RO11		In the last six months, someone at work has talked to me about my progress.	
RO12		This last year, I have had opportunities to learn and grow.	
2	Contribution	How do you think your experience as an employee of [Org] compares to other employment opportunities?[11]	Close-ended (likert scale)
RO1		A lot worse	
RO2		Worse	
RO3		No change	
RO4		A little better	
RO5		A lot better	
RO6		Don't know	
3		What could [Org] do to improve as an employer?	Open-ended
RO1		Write in	
4	Risk	Did anything unexpected - good or bad - happen because of [Org]? Please provide details.	Open-ended
RO1		Write in	

[10] The Gallup 12 questions were sourced at: https://www.goalbusters.net/uploads/2/2/0/4/22040464/gallup_q12.pdf
[11] This question provides an employee perspective on how their employer compares to others within the same sector. These results can be checked against industry standards across other non-self-reported indicators such as salary, benefits, working conditions and attrition rate.

Conclusion

Conclusion

For all the enterprises surveyed, the self-reported data is complemented by non-self-reported data. For examples of this see Appendix 1 (p. 47-48).

Feedback from the enterprises suggest the questions were useful for impact management

- Feedback received from the enterprises' management teams included:
  - "The sector often reports change in income as the most important outcome, so we appreciated that the survey captured other outcomes related to quality of life and asked respondents about their perceived importance."
  - "When we first saw the survey, we thought the results would be generic – as few questions were customised to the company… But the results have been actionable and we have already implemented a working group to create solutions in response to the feedback."
- The management teams interviewed after the pilot said that they would be keen to use the questions again. A few said that they would be likely to make some adaptations and modifications to help them tease out more specific insights relevant to their business model, and follow up to see whether changes made as a result of the initial survey had been effective.

Self-reported data is part of, but not the whole solution, and implementation should be iterative, no matter what questions are asked

- Surveys - as with any method to collect data on impact performance - are best executed iteratively, with one data collection exercise informing the next.
- This pilot has shown that a good way to begin impact measurement is to start with a survey. The self-reported data gathered through a survey is useful to both inform the design and marketing of a product or service, but also as a low-cost initial exercise to determine what additional impact measurement methods might be required to fill any gaps.
- Self-reported data is especially valuable for assessing the What dimension, as it is a reliable method for uncovering other positive and negative outcomes that the enterprise itself may not have intended to deliver. In addition, collecting direct feedback from customers on which of these outcomes is most important to them helps management teams prioritise which impacts they manage.
- The value of self-reported data also varies according to the type of product or service, the market, and the availability of the same information in other ways. The resources for measurement activities are often very limited, so measurement system design should prioritise the most reliable approach to evidence generation that is proportional in cost to the decision it drives.
- The application of a standard question set should be filtered through a quick assessment of what other data is available. However, where good data already exists for some or all of the dimensions of impact, surveys might still provide insights on unexpected negative impacts, helping the enterprise manage their risk.

Self-reported data may be simple, but collecting it takes care.

Just about any feedback is useful, even from a handful of customers. But as with any data collection exercise, in order to have trust in the answers (and therefore confidence in any decisions) considerations of question robustness, bias, and sample size significance are key.

A few especially important implementation considerations are:

Question robustness

- Feedback is inherently subjective and individual views can change. Whilst this is to be expected (humans are inherently adaptive to both positive and negative change), a good feedback question should not change dramatically over short period of time (unless there is a genuine external event or shock that drives the change.) In other words, you want a question to avoid the “got out on the wrong side of bed” effect, whereby you might receive significantly altered responses depending on mood.
- To assess the performance of our questions, we back-checked responses to certain questions. Researchers call back a sample of customers within a few weeks of the initial survey and re-ask several survey questions. In our initial back checking calls, most questions performed well, but we have found variations for the questions around time of onset of quality of life changes, and level of importance of these changes. Understanding why responses vary will require further investigation. As you use these questions, or add feedback questions of your own, we encourage regular back-checking.

Bias

- There are many potential causes of bias, the tendency of measurement to give an over- or under-estimation. Two interlinked causes to pay particular attention to with respect to feedback are low response rates and/or systematic differences in the representativeness of your respondents.
- Consider the possibility for self-selection bias, i.e. only people with extreme views responding to your surveys. This can be managed with precise statistical techniques such as re-weighting, but in most cases, this may be beyond the implementer or a disproportionate expense. In such cases targeting a generically high response rate, say greater than 50%, will mean that any self-selection of the most opinionated respondents is diluted. Additionally, interrogation of the data itself may give clues: do you see a healthy spread of responses, or only extremes?

Significance

- Typically, this is a matter of judgement. Decisions need not be made on a 95% significance level. We believe more attention should be paid to the bias and question robustness. Some data is typically better than none, and a representative sample of 200-300 is typically ample for most feedback surveys.

Stakeholder engagement is not a one-time exercise

Ideally, stakeholders are involved throughout an enterprise’s decision-making process. Visit Social Value International[12], Keystone and 60 Decibels’ websites to read more on how to ensure stakeholder perspectives are considered throughout an enterprise’s decision-making processes, along with further guidance on best practices to abide by when engaging stakeholders.

[12] Social Value International’s standard on involving stakeholders can be found here https://socialvalueint.org/social-value/standards-and-guidance/standard-on-applying-principle-1-involve-stakeholders-2/

Appendix 1: Using both self-reported and non-self-reported data

Example 1. This table summarises data on gym users who reported positive changes as a result of their gym membership, illustrating how it is often useful to examine non-self-reported data and self-reported data together. The same exercise should then be carried out to assess the impact on those reporting less positive changes (e.g. the 10% seeing no improvement) and those experiencing negative changes.

EXAMPLE 1 An illustrative gym effect

			Self-reported data		Non-self-reported data		
Dimension		**Impact Category**	**Question**	**Result**	**Indicator**	**Value**	**Assessment**
What		Outcome level in period	When you first visited [Org] what improvement were you looking for?	"Improved fitness"	Average gym visits per week (proxy for increased fitness) Active members	1.3 78%	Positive
What		Threshold for positive outcome	Is the improvement you are experiencing sufficient to meet your needs?	"Yes" or "Maybe" for 89% of respondents	Recommended hours of exercise per week[13]	2	
What		Importance of outcome to stakeholder	How important are these changes to you?	Very important for 45% of respondents	N/a	-	Important
What		SDG				SDG 3.4	
Who		Stakeholder			N/a	Customer	
Who		Geographical Boundary			N/a	Spain	
Who		Baseline	First time gym users	45%			Underserved
Who		Stakeholder characteristics					
How Much	Scale	Number of <stakeholder> experiencing <outcome>	Number of respondents citing improved fitness	50	Total gym members	106,000	Medium scale
How Much	Depth	Degree of change experienced by <stakeholder> as a result of <effect>	To what degree did you experience these improvements?	"Improved a lot" (54%) "Improved a little" (35%)			Deep
How Much	Duration	Time period for which <stakeholder> experiences <outcome>	Have these improvements been long-lasting?	"Long-Lasting" (40%) "Too soon to know" (53%)	Average tenure for gym users	14 months	Medium term
How Much	Rate	Time taken before <stakeholder> experiences <outcome>	How quickly did the improvements happen?	'Within weeks' or 'within months' for 69% of respondents			
Contribution	Depth	Estimated degree of change that would occur otherwise for <stakeholder>	Is there a good alternative to [Org] that will deliver the life improvements you want?	"Yes - the same" (22%) "Yes, but this company is better" (17%)	Average tenure for first time gym users nationally	9 months	Likely better
Contribution	Duration	Estimated time period that <outcome> would last for otherwise					
Risk		Evidence risk				Data available across most dimensions	Low risk
Risk		Stakeholder participation risk	Customers surveyed	Low risk			
Risk		Endurance risk	Do you intend to sign up with [Org] again?	Yes (48%) "Maybe" (33%)			Low risk

[13] https://www.hsph.harvard.edu/nutritionsource/2013/11/20/physical-activity-guidelines-how-much-exercise-do-you-need/

Example 2. This table illustrates how self-reported data can be used alongside non-self-reported data to uncover different outcomes. In this example weight-loss progamme participants who reported positive changes in quality of life, separate from their healthy weight outcomes.

EXAMPLE 2	An illustrative weight loss effect		Feedback		Objective Measurement		Assessment
Dimension	**Impact category**	**Commentary**	**Question**	**Result**	**Indicator**	**Value in period**	
What	Outcome level in period	The outcome experienced by the <stakeholder> when engaging with the enterprise.	What changes have you experienced?	"Quality of life improvements" (60% of respondents)	Current period body mass index (avg)	24	Positive
	Threshold for positive outcome	The level of outcome that the <stakeholder> considers to be positive/ good enough.	How would you rate your current quality of life from 1-10?	8 or higher (90%)	Healthy body mass index (avg)	23	
	Importance of <outcome> to stakeholder	Stakeholder's view of whether the outcome they experience is important.	How important are these change to you?	Very important (100%)		-	Important
	SDG					SDG 3.4	
Who	Stakeholder	The type of stakeholder experiencing the <outcome>.			N/A	Customers	
	Geographical boundary	The geographical location of the <stakeholder>.			N/A	Dallas, USA	
	Baseline	The level of <outcome> being experienced by the <stakeholder> prior to engaging with the enterprise, i.e. baseline.	Baseline result: how would you rate your current quality of life from 1-10?	6	Body mass index prior to programme (avg)	27	Underserved
	Stakeholder characteristics	Relevant stakeholder characteristics	Have you tried to lose weight in the past?	65% 'yes'	% with chronic disease	50%	
How much: Scale	Number of <stakeholder> experiencing <outcome>	The number of individuals experiencing the <outcome>. When the planet is the <stakeholder>, this category is not relevant.	Number surveyed	233	Total # programme participants	450	Large scale
Depth	Degree of change experienced by <stakeholder> as a result of <effect>	Degree of change is calculated by comparing outcome for stakeholder prior to engagement with enterprise to outcome value in the current period.	Delta change on quality of life rating change	Average of 3 point improvement	Average BMI point improvement	3 points	Deep
Duration	Time period for which <stakeholder> experiences <outcome>		How long-lasting were these changes?	1 year or longer (60%) 6 months or longer (40%)	Duration for which participants weight remained below 24 BMI	50% - 2 years 40% - 3+ years 5% - unknown 5% <1 year	Medium-term
Contribution: Depth	Estimated degree of change that would occur otherwise for <stakeholder>	Estimated degree of change likely to occur otherwise for the <stakeholder>.	Is there another service available to you that could have delivered these changes?	30% - yes 30% - no 40% - don't know	Census data: in region, % change in overweight people in period	+5%	Likely better
Duration	Estimated time period that <outcome> would last for otherwise						
Risk	Evidence risk				N/A	Low (data available across other dimensions)	Low risk
	Stakeholder participation risk		N/A	Low			
	Unexpected impact risk		Have any unexpected changes occurred?	Non (38%), positive (60%), negative (2%)	% participants experiencing no improvement	1%	
	Endurance risk		Did you finish the programme?	90% yes	% course attrition rate before completion	5%	

Appendix 2: Technological differences to be aware of for implementation

The chart below highlights some of the major pros and cons of varying data collection modes. For more detail we recommend the Lean Data Field Guide.

TABLE 2 Technological differences[14]

Technology	Pros	Cons
SMS	+ Ability to incentivize customer with airtime bonus	– Need high mobile penetration & decent literacy rates
	+ Customer can complete at his/her convenience	
	+ Good quality open-ended questions	
	+ Honest responses to sensitive questions	
	+ Ability to compare responses to general population panel (requires existing panel - e.g. Kenya)	
IVR	+ Works in low-literacy areas (e.g. parts of rural India)	– Multiple choice only, not able to analyse qualitative feedback easily
	+ Honest responses to sensitive questions	– Less customer-control over timing of survey
Phone Calls	+ Works great in low literacy contexts	– Still need phone numbers
	+ Good for more complicated questions that require explanation/probing ("tell me more about that")	– 3-5x more expensive than IVR[13] & SMS
	+ Higher confidence in the quality of data	
In Person	+ Overcomes low mobile phone penetration	– Time consuming, expensive (cost depends on quality of enumerators)
	+ Allows observation (e.g. see household environment to witness how a product is being used	– Experience can be intrusive to customer
	+ High quality, trusted data	

[c]Source: The Lean Data field guide, https://acumen.org/wp-content/uploads/2015/11/Lean-Data-Field-Guide.pdf

[14] Interactive Voice Response (IVR) is an automated telephony system that interacts with callers, gathers information and routes calls to the appropriate recipients.

Acknowledgements

About 60 Decibels

60 Decibels helps social change organizations listen to the people that matter most - their customers, beneficiaries, suppliers or employees – in order to make more positive impact in the world.

Initially created as Lean Data, by the not-for-profit Acumen, to meet its own social impact measurement needs, 60 Decibels is focused on providing high-speed, high-quality, end-to-end customer insights and impact measurement services, anywhere in the world.

Working for some of the sectors best known investors, 60 Decibels is also building some of impact investing's deepest benchmarks of social impact, allowing organizations to see how they're performing relative to their peers for the first time.

About Keystone Accountability

Since its founding in 2004, Keystone has been developing practical and sustainable methodologies for impact measurement.

After several years of intensive work in the fields of impact evaluation, performance management, standards accountability and certification, and knowledge management, Keystone developed a distinctive method for asking questions, analysing the answers, and then deliberating on these findings, which it calls Constituent Voice™ (CV).

Over the past decade Keystone has worked with over 185 clients and partners in 17 countries to introduce and adapt CV as an alternative approach to performance and impact management. CV works best where the quality of relationships between an organization and its constituents (e.g., clients, beneficiaries, users, etc.) matters for realizing impact.

Consulting input was also provided by Social Value International.

We would like to thank all the participating enterprises and our partners at Acumen and Bridges Fund Management for their willingness to pilot these surveys and share the findings for the benefit of the wider market.

IMPACT
MANAGEMENT
PROJECT

Impact Management Project, May 2019