PREVIEW BEGINS **OCT 2014** LAUNCHES **JAN 2015**

INTRODUCING WWW.FEEDBACKCOMMONS.ORG

A serviced cloud platform to 'share and compare' feedback data and collaboratively develop feedback and learning practice

CONSTITUENT VOICE

TECHNICAL NOTE 1

Version 1.1, September 2014

"At its heart, Keystone's Constituent Voice methodology promises to shift the very power to define 'success', and to declare when it has been achieved, into the hands of those that development organizations claim to serve. This shift toward greater agency is essential if we are to make development more effective."

Raymond C. Offenheiser • President, Oxfam America

keystone

www.KeystoneAccountability.org

Table of Contents

How to read this technical note

Constituent Voice™ is a methodology developed by **Keystone Accountability** to enable organizations to improve results by optimizing their relationships with their constituents. The purpose of this note is to help organizations understand what Consituent Voice method is and how it works.

Anyone can use the Constituent Voice methodology for free. In order to provide you with the emerging learning and experience relating to Constituent Voice and feedback systems, we invite you to register your interest at **www.keystoneaccountability.org**

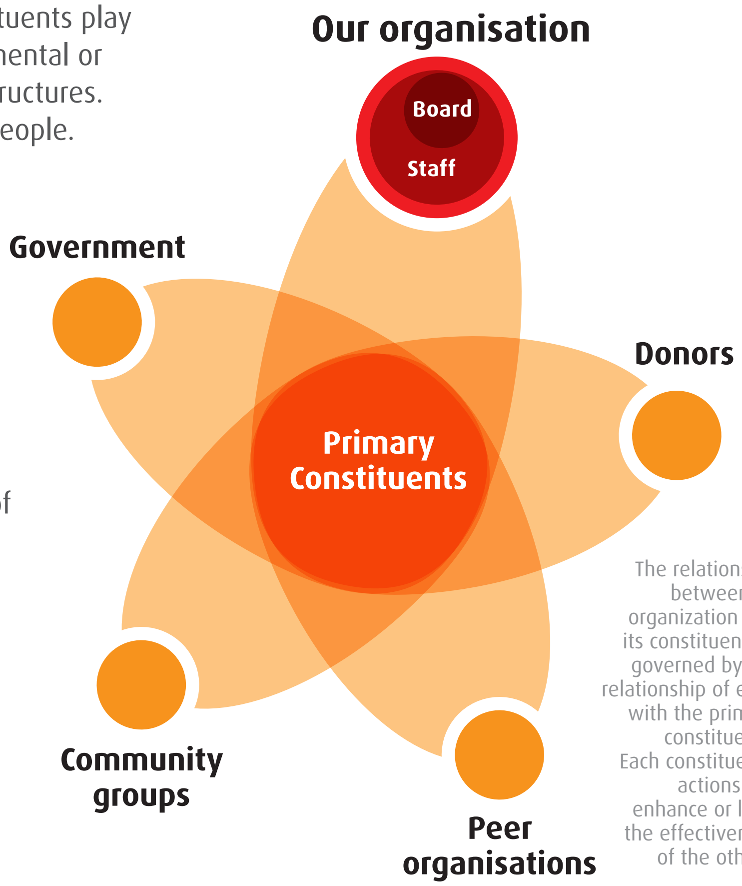
Constituent Voice is a work in progress. We have a strong focus on rapid cycle learning across all our work, including and maybe even especially with respect to our core methodology. We are constantly searching for better questions and new ways to ask then, for new approaches to analyzing and interpreting data, and for simple ways for organizations to respond and improve. While comprehensive in the sense of describing all aspects of our methodology, this note is not exhaustive in its illustrations. Examples of Constituent Voice in action are recorded and published on our **website** continuously, and we will update this Technical Note regularly.

Organizations have a diverse range of constituents in addition to the people who are the primary focus of the organization's mission.

Some constituents contribute directly to, and are directly affected by, the organization's programs and activities. These include staff, donors, and partner organizations. Other constituents play a less direct role, like board members, governmental or traditional authorities, and wider community structures. Sometimes an organization's constituents are people. Sometimes they are other organizations. Nevertheless, the organization usually has a degree of moral or legal obligation to listen and engage with each of its constituent groups.

Keystone helps clients to cultivate voice with all these constituents.

This note is written mainly about applications of CV where respondents are individuals. To learn more about how we cultivate organization-to-organization feedback, please visit the Performance Surveys pages on the Keystone **website**. There you will learn how some of our clients are using CV to improve their relationships with investees, suppliers, grantees, and partners. Our data set here includes over 4,000 responses to Keystone Performance Surveys from over 30 countries.



The relationship between an organization and its constituents is governed by the relationship of each with the primary constituents. Each constituent's actions can enhance or limit the effectiveness of the others.

Underpinnings

Constituent Voice is a blend of two important lines of thought in development.

Constituent Voice's primary theoretical influence comes from Amartya Sen and Jean Drèze, who locate individual human agency at the heart of development. Sen's great synthesis work, *Development as Freedom*, concludes that development cannot be reduced to material well-being as evidenced by basic incomes, nor be achieved by rising average per capita incomes. Rather, it requires a package of overlapping political and economic mechanisms that progressively enable the exercise of a growing range of freedoms that allow people to meet their basic needs and unlock their innate abilities for self-determination. When people express their freedom, we call it Agency.

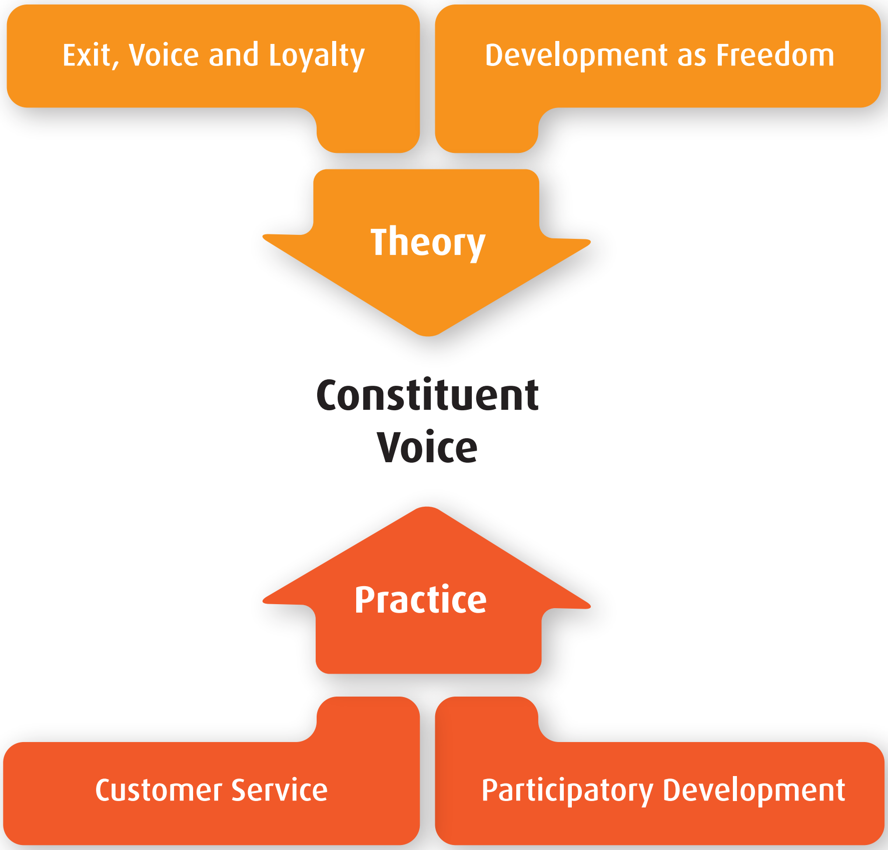
Constituent Voice's other main intellectual debt is to the seminal work of Albert O. Hirschman on the nature of choice under limited choice conditions.[1] Hirschman observed that when faced with unsatisfactory performance from an organization, people might decide not to exit but to ante up – to engage to improve the organization. Hirschman called this Voice.

> “... when faced with unsatisfactory performance from an organization, people might decide not to exit but to ante up – to engage to improve the organization ...”

Our metrics track Agency and Voice (among other things) and show them to be essential assets for development service providers. Constituent Voice method has discovered that feedback loops between constituents in development can accurately measure and cultivate these all-important personal and framing conditions.

Because Agency and Voice are so central to who we are as sentient social beings, they are both means and ends of our social, economic and political expressions.

1 We have named the Constituent Voice operational cycle after Hirschman. For a further elaboration of our intellectual debt to Hirschman, see our **Feedback Labs** blog, “Voice As Both Means And End”. **http://feedbacklabs.org/voice-as-both-means-and-end/**



Just as Constituent Voice blends two intellectual traditions, it also blends two practice traditions. Both have been widely utilized for over 60 years but have never before been combined. From the world of development, Constituent Voice draws on a succession of *participatory development techniques* that emerged out of the field of action research in the 1950s – such as rapid rural appraisal, participatory evaluation, appreciative enquiry and many others. From the world of consumer facing businesses, Constituent Voice draws on the customer service industry. Customer service metrics trace their roots to the consumer rights social movements, which not coincidentally also emerged in the 1950s.

One headline objective of Constituent Voice derives directly from the success of the *customer service field*. We hypothesize that just as customer service metrics have proven to be reliable predictors of business success, Constituent Voice metrics can predict development outcomes. In the world of development and social change, we have never done the work of systematically tracking relationship metrics and comparing them to outcome metrics. Constituent Voice is now doing this work.

No shortcuts to progress

Over the past few years of working with diverse organizations, Keystone has learned that there are no shortcuts to the successful cultivation of Constituent Voice. Every step in the operational cycle of Constituent Voice, which we have termed the Hirschman Voice Cycle (pictured below), is necessary to realize intended outcomes.

Our overriding lesson from the past few years of client work is that collecting feedback – listening – is not enough. It is necessary to land what you hear in transparent performance metrics. In order to understand the stories behind the measures, it is necessary to hold a dialogue with your constituents. We call this sense-making. Then you must act on the resulting insight. Finally, ongoing Constituent Voice feedback will let you know if corrective actions are having their desired effects.

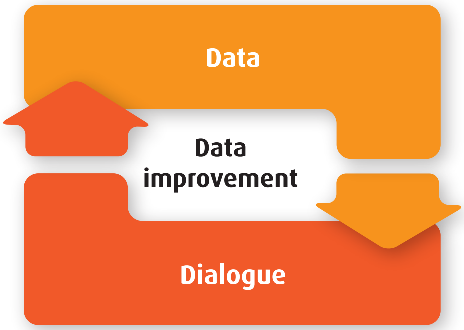
This overall observation about Constituent Voice is strongly reinforced for us by three common errors that we have seen in organizations that fail in their efforts to improve through feedback from their constituents.

The first and most common error is to think that by collecting feedback you have enabled Voice. A broken chain pulls no weight. The most likely fate of collected evaluation and monitoring data in development today, regrettably, is to lie under-used – neither analysed nor properly understood, never reported back to clients or service users, and rarely acted upon.

> … too much evaluative practice in organizations is confined to a kind of internal ghetto, or gulag, that operates on its own cycles that are too long to be effectively integrated into day-to-day management …

The second common error is to over-invest in feedback data collection with tools designed for social research rather than for performance management. This error typically involves going to the field with long surveys that are costly to implement and burdensome for respondents. Moreover, such tools produce overwhelming volumes of data that staff relegate unused to drawers, boxes and computer folders.

Constituent Voice takes a management perspective on data collection, that "even imperfect data can save lives".[2] Once Constituent Voice is understood as an ongoing and continuous process, the stakes with respect to a particular instance of feedback data change considerably. The initial data collection does not have to be perfect as it is subsequently refined and validated through dialogue, insight and improved relationships.



The third error might be described as "ghettoization" of evaluative activity. Too much evaluative practice in organizations is confined to a kind of internal ghetto, or gulag, that operates on its own cycles that are too long to be effectively integrated into day-to-day management. These gulags tend to be led by specialists and experts who have no line authority over frontline workers and who tend to go outside the organization to undertake research and evaluation activities. Constituent Voice addresses this by embedding itself directly in normal management practice. As is explained below, frontline workers are expected to be creative evidence-driven problem solvers through two-way communications with constituents.

The discussions in this report of each of the steps of the Hirschman Voice Cycle indicate ways to avoid these common errors.

2 "Even Imperfect Data Can Save Lives, Suprotik Basu, 9 September 2013. **http://www.impatientoptimists.org/Posts/2013/09/Progress-on-MDGs-Requires-an-Obsession-for-Tracking-Results**

Enabling conditions and readiness

Before introducing Constituent Voice surveys, it is necessary to establish the minimum conditions for success. The very idea of Constituent Voice is new for most people. It takes some getting used to. So the first step to introducing Constituent Voice at any organization is to communicate the value and benefit of using the Constituent Voice cycle in a way that allows people to understand and even get excited about it.

The best ways to do this vary according to the constituent group and senior management need to craft an approach for internal and external constituents. Frontline employees often worry about the possibility of negative feedback about their performance. Board may worry that negative feedback will undermine donor relationships. These and other natural fears need to be addressed.

> “... only when they understand how it is in their interest to participate will you earn the kind of survey response rates that you want. ...”

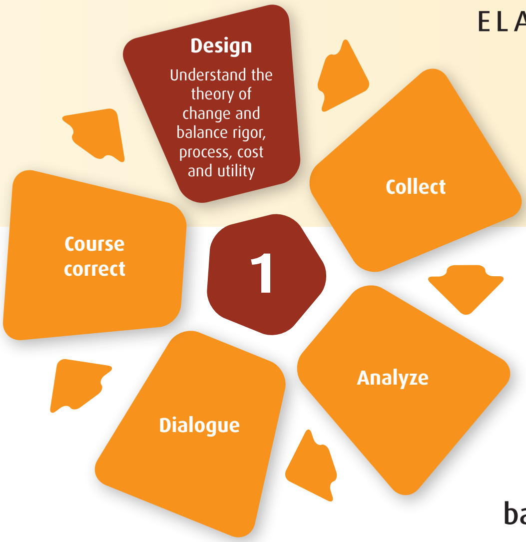
Similarly, the constituents who are going to be providing feedback don’t have a context at first in which to see the value in providing feedback. Only when they understand how it is in their interest to participate will you earn the kind of survey response rates that you want. In using non-anonymous feedback, extra care needs to be used when explaining who will see the data, for what purpose, and in what form. This theme will recur as we discuss data collection and also dialogue with feedback providers.

Another dimension to getting started with Constituent Voice is to understand your organizational capacity in relation to the demands of doing Constituent Voice. While Constituent Voice is not objectively hard to do, our work with organizations over the past few years suggests that doing Constituent Voice does drive organizational change. It is extremely helpful to get on the front foot of this organizational change process by building a list of the kinds of activities that will become second nature through the introduction of a Constituent Voice system. For this purpose, we recommend that organizations use the readiness diagnostic tool operated by Feedback Labs.[1]

1 Feedback Labs is a consortium of organizations, including Keystone, committed to making governments, NGOs and donors more responsive to the needs of their constituents http://feedbacklabs.org/.

Testing survey questions to ensure their relevance and salience for women in Sindh, Pakistan

Stage 1 • Design



Balancing four measurement design principles

Design is critical. Constituent Voice design seeks to strike a balance across four central principles of good developmental measurement practice:

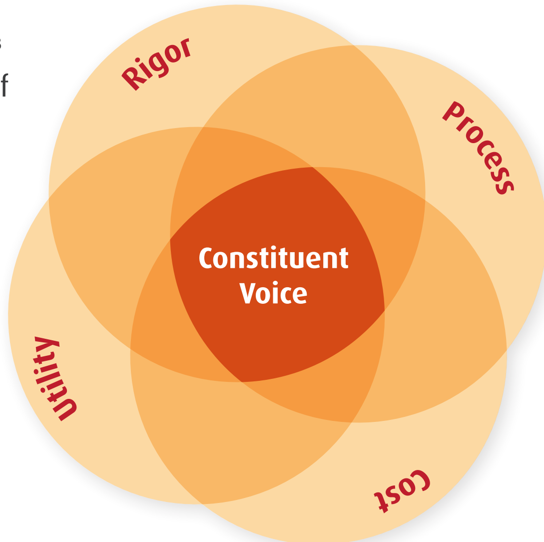
1. **Rigor** – The data we use to form judgements and make decisions should be accurate. We apply good social research and evaluation practices to address common biases in feedback data, including those that arise from the various methods we use to collect feedback. These include selection bias, various sample-related biases, and interpretation bias associated with the unconscious cognitive patterns of those designing and interpreting surveys (or focus groups, semi-structured interviews and observational studies for that matter).

2. **Sensitivity to process and culture** – Development at its best grows the capacity of people to discover solutions and take control of their lives. It is a living, generative process, and this means we need to measure and nurture those things that give life to an intervention: attitudes, relationships, capabilities, and agency. These are stepping-stones to the changes in material conditions that development aims for. Constituents must feel that they belong and contribute. Constituent Voice asks questions in a way that is respectful of constituents, for purposes that they understand and endorse. Constituent Voice systems are designed to build trust and confidence between an organization and its constituents, to generate knowledge and insight that constituents can and do use for ends that they set, and, in the process, to grow their capacity.

3. **Cost** – If resources are infinite, then anything can be measured precisely. Resources are finite, however, and organizations must make difficult choices between money spent on measurement-related activities and money spent on the intervention itself. The Constituent Voice cost/value proposition is that for less money than conventional monitoring and evaluation, Constituent Voice will deliver more value in terms of contribution to intended outcomes. One cost advantage of Constituent Voice is the ability to laser in on specific problems, generating ideas for solutions, and signalling the effects of corrective actions. Another advantage is that the doing of Constituent Voice has the tendency to contribute positively to the intervention. It is not designed to be a neutral measurement exercise that has no bearing on the thing being measured. Constituent Voice is designed to enable the organization and its constituents to be more mindful of what is happening and not happening, and to adjust their behaviors to improve results.

4 **Utility** – The primary measure of evaluation has to be its utility. Does it lead to improvements? Constituent Voice is designed to generate data that is useful for the key constituents of an intervention or organization – including the staff and board of the organization, the people who are meant to enjoy the value being created, the financiers, the wider field in which the work is located, and society at large. But first and foremost, Constituent Voice is designed to be useful to the organization undertaking the work. It is intended to significantly enhance relationships with constituents through more authentic conversations and accelerate outcome attainment. Constituent Voice is a system for continuous improvement in relationship quality and general performance. It is has a real-time orientation and features short measurement-reflection-action cycles. This is implicitly and explicitly part of the Hirschman Voice Cycle.

Finding the right balance across these four variables is difficult to achieve. In recent times, one principle – statistical rigor – has been elevated above others as some kind of 'gold standard'. Gradually, a more sophisticated consensus is emerging that sees use as the paramount purpose.[3] More practical materials are being produced to support the use of evaluative findings.[4]



To strike a balance that is right for a particular intervention, we start with a careful enquiry into the context and the theory of change of the intervention. From there we propose and test methods and instruments to collect feedback. It is essential to involve those who are to be providing feedback at this point in the design process. We look for:

- Mechanisms of collecting feedback that are affordable, convenient to the organization, and can extend to large respondent groups;
- Mechanisms of collecting feedback that are unobjectionable and non-burdensome to the respondents;
- Questions understood in the same way across the respondent group;
- Questions using Likert scales that can be used over time and across units or organizations to enable comparison and sense-making dialogues; and
- Getting the right amount of data – enough to enlighten and spur further investigation, not enough to drown us.

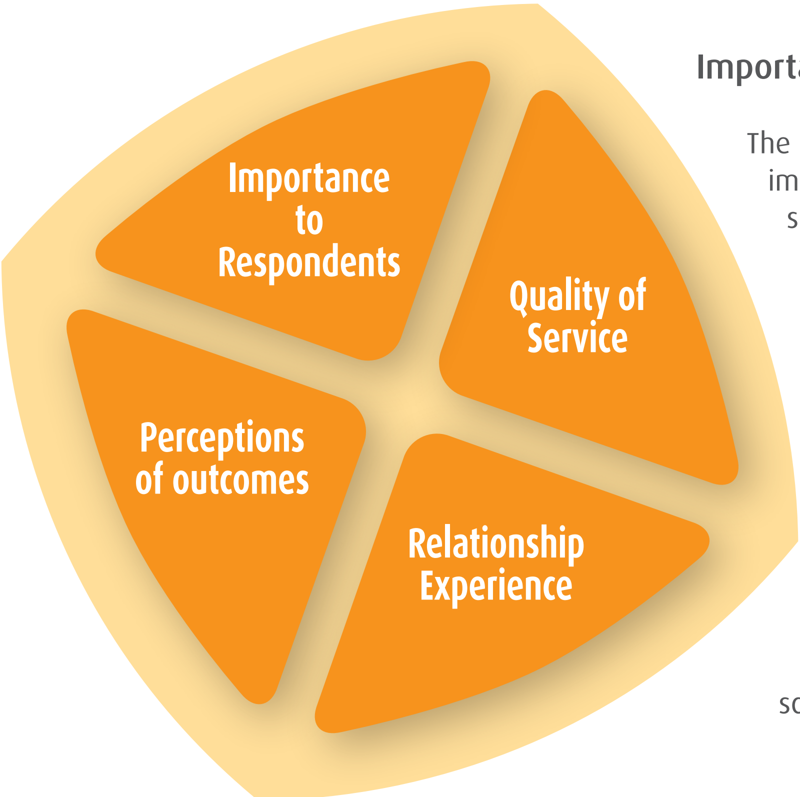
3 For a blog making this point unequivocally see "Time for a Gold Standard of Use", Fay Twersky, 27 February 2012. http://www.effectivephilanthropy.org/blog/2012/02/time-for-a-gold-standard-of-use/

4 InterAction in late 2012 produced a four-part guidance note series on impact evaluation that concluded with a note written by Keystone's David Bonbright entitled, Use of Impact Evaluation Results. http://www.interaction.org/document/guidance-note-4-use-impact-evaluation-results

Evidence on four dimensions of performance

Our work over the past few years helping organizations to cultivate Constituent Voice has shown that Constituent Voice survey methods can collect reliable evidence of performance across four key and interrelated dimensions of organization performance:

It is necessary to create a blend of questions across the four categories to get the full picture of the respondent's experience of and attitudes toward the organization or intervention. The relative utility of each category varies for each case depending on the nature of the work, the context, and the theory of change.



Importance

The importance category seeks to establish the importance or relevance of the organization, service or product to the respondent. Is it central to the respondent's life or goals? Or is it relatively trivial? Is it somewhere in between? Typical questions include:

- How important are the services provided by [...] to you and others in your area?
- How important is this issue [or cause] to you?

Importance scoring is also used within an organization, as exemplified in the ranking schematic opposite:

Please describe the most common 5 services that you have received from [...] over the last 12 months. Then rank each activity in terms of how important it is to you.

Activity	Importance rating										
	Not important					< >					Extremely important
	0 ■	1 ■	2 ■	3 ■	4 ■	5 ■	6 ■	7 ■	8 ■	9 ■	10 ■
	0 ■	1 ■	2 ■	3 ■	4 ■	5 ■	6 ■	7 ■	8 ■	9 ■	10 ■
	0 ■	1 ■	2 ■	3 ■	4 ■	5 ■	6 ■	7 ■	8 ■	9 ■	10 ■
	0 ■	1 ■	2 ■	3 ■	4 ■	5 ■	6 ■	7 ■	8 ■	9 ■	10 ■
	0 ■	1 ■	2 ■	3 ■	4 ■	5 ■	6 ■	7 ■	8 ■	9 ■	10 ■

Quality of Service

Questions here hone in on service quality, ideally as close as possible to a unique touch point, such as a training course, a sales exchange, or a counselling session. Questions here are both general and specific.

Examples of specific questions from actual Keystone surveys include:

- I get the attention and support I need from [...] staff.
- We keep busy and learn something in this class every day.
- Based on the services I receive from [...], I would recommend [...] to family or friends.
- Have you received this same service from another organisation? If "Yes", how does it compare, much worse, or much better?
- Did the service happen at a place and time that was easy for you to attend?
- Did they have the necessary resources (teaching materials, equipment, facilities etc.) to achieve the purpose?
- Were you able to participate effectively?

Quality of Relationship

Relationships matter. Precisely how they matter varies according to the nature of the intervention. Our relationship questions enable an organization to determine precisely what their relationship quality is, and how to improve it. Overtime, they reveal how relationship quality affects effectiveness and outcomes.

Our work over the past few years has discovered some core elements of relationship quality for most organizations, notably trust, fairness, voice, and empowerment. These relationship building blocks are interrelated and overlapping. By asking related questions, some focussed on specific interactions and some more generally about the relationship overall, we can compare answers and build up a nuanced picture of the organization's constituent relationships.

Because there is more than one side to every relationship, we also routinely ask respondents to rate their own readiness to engage the organization.

Common relationship questions used in our surveys include:

Trust (incorporates confidence, integrity, manner, credibility, professional ethics)

- I have confidence and trust in the integrity of [...].
- I have confidence in the technical skills of [...].
- Does [...] fulfil its promises - do what it says it will do?
- Do you feel that workers from [...] really care about you and want to help you as best they can?
- Does [...] have the necessary knowledge and skills to do it well?
- I feel that [...] is working sincerely and honestly for my benefit.

Fairness (incorporates pride and respect)

- [...] treats me fairly.
- Staff at [...] always treat me fairly.
- People at [...] treat me with courtesy, dignity, and respect.
- I am proud to be associated with [...].

Voice (incorporates responsiveness)

- It is worth my effort to engage with [...] to try to get it to do what I think is important.
- I feel completely free to ask questions and say what I really think.
- I have the opportunities I need to tell [...] what I think about its work.
- I feel that I contribute meaningfully to [...].
- I believe that [...] will satisfactorily respond to and act upon my feedback.
- I am sure that [...] will use my answers to this survey to improve its services.
- How much do you influence the choice of services offered by [...]?
- [...] listens to my questions and my views and responds in a genuine way.

Empowerment (incorporates agency and inclusion)

- Because of [...], I am more able to stand on your own feet and achieve what I want.
- Is [...] helping you to stand on your own feet, take your own decisions, find your own solutions to problems on your own?
- With support from [...], I am getting better at achieving my goals.
- Because of [...], I have more positive relationships with other people and organizations that can help my life improve.
- I am more connected to the community and community resources thanks to [...].

Readiness (promoting self-reflection by the respondent)

- I feel that I am ready and willing to try new things offered by [...].
- I know that the more I put into [...], the more benefits I will get.

Outcomes

Constituents have a direct and often dispositive experience of outcomes (or their absence). Constituent perceptions of outcomes are evidence of outcomes that we like to combine with other measures of outcomes to arrive at a composite outcome narrative. This is discussed further under "Analysis: Triangulation".

- With help from [...], I am making progress on my goals.
- I have more income and my family lives a better life because of [...].
- There are real and beneficial changes taking place in my life and my community because of [...].
- Are you doing anything differently after this receiving the service? Yes/no
  - If yes, what? __________
- How well did the [advice/technology/methods/service] work when you tried them out? [Response option includes, "I have not tried them out"]

Stage 2 • Collect

The three most important things to consider in feedback data collection are context, context and context.

Conventional social science-based surveys reckon with the complexities of context by experimental methods such as randomization across controlled comparator groups. If you are prepared to dedicate a large portion of your entire budget, this approach can tell you what effects you are actually causing. But it will not generate continuous, real-time data that you can embed in performance management systems to improve relationships and accelerate intended outcomes.

Constituent Voice tackles the complexities of context by locating the process of data collection as one step in a five-step cycle that over time generates higher rates of participation in surveys and ever more honest and candid feedback from respondents.

Those who practice Constituent Voice understand that they must earn their constituents' participation and candid feedback. In order to get great feedback quality, they must demonstrate to their constituents that good things are most likely to happen when respondents are candid and frank in their feedback; in other words, that the effort to provide feedback is rewarded by the improvements that follow. It is about creating a culture in which formalized systematic feedback is understood by all constituents – from staff to society-at-large – as a means to more authentic conversations about what is working and what is not, how to improve, and what different constituents must do in order to realize that improvement. This culture does not happen overnight, and can take time to cultivate. But once it is there it endures through staff turnover and other forces of change.

> “... those who practice Constituent Voice understand that they must earn their constituents' participation and candid feedback....”

Accordingly, the first step in Constituent Voice data collection is to explain its purpose with intended respondents. This explanation typically includes the following elements:

- A statement of the purpose of the survey and how it fits into a larger purpose of ongoing dialogue and continuous improvement.
- An outline of and timeline for the steps that will follow from data collection, including reporting back the findings of the survey/data collection method to respondents, and a commitment to keep surveying in order to see if resulting corrective measures are working.

- Some examples of the kinds of things that the organization hopes will happen as a result of the survey. This helps respondents get a tangible sense of the possible benefits of their participation.
- Any explanations that may be required for the more technical characteristics of data collection. For example, in the case of a survey, whether the survey responses are anonymous (and how that anonymity is protected).
- An ethical code and grievance procedure.

These framing conditions for a Constituent Voice system are repeated continuously as part of a long-term culture building and affirming process.

Anonymous vs. non-anonymous & independently collected vs. self-collected

From the organization's point of view, non-anonymous data is far more useful than anonymous data as it can be matched to other data the organization has relating to the respondent. This helps in interpreting feedback, as well as identifying correlations between feedback and later outcomes. While it is possible to develop some degree of conviction on these things by using characteristics of respondents that do not include their identity (such as age, gender, length of relationship, locations, etc.), non-anonymous responses mean that you do not have to ask these demographic questions in the survey (saving everyone's time) and of course the analysis is far more precise.

But from the respondent's point of view, "on the record" responses can be problematic for different reasons ranging from uncertainty about the possible negative consequences to power dynamics at play in the relationship. For this reason, anonymous feedback is generally less biased than non-anonymous feedback, and

I'm from the company and we want your feedback

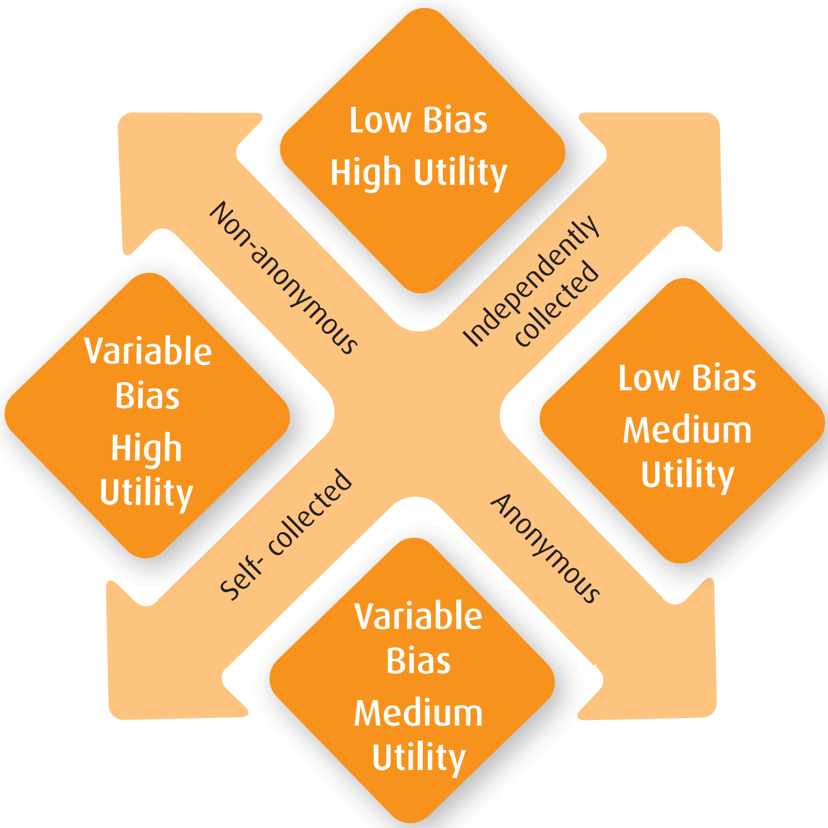
In order to test courtesy bias, we ran two versions of the same household survey simultaneously across 12 villages in Tanzania. Respondents were selected randomly. The client conducted one survey and the other survey was conducted by Keystone. The same introduction to the survey was given promising anonymity of responses. Respondents were aware, however, when they were taking a survey directly from the client. The findings from this experiment were striking. For every question in which one could expect to see a courtesy bias (for example, rating the fairness of the client), there was 30 percent more positive response from the client-administered survey. For more factual questions or questions where one would not expect to see a bias (such as what benefits the respondent received from the company) the answers in the two surveys were indistinguishable.

at least initially is the preferred choice. We say initially, as when an organization has been gathering feedback for some time, and there is an open culture of honesty, it is possible to revert to non-anonymous feedback.

Similarly, when an independent party with specialized skills collects the data, it is more likely the process will reduce response biases. From the organization's point of view, on the other hand, collecting the data themselves may be easier, less costly, and more quickly embedded in normal performance management.

These dynamics are visualized in the figure below. In it, we assume that bias and utility are independent variables, though of course in practice they are not. The essential take-away here is captured in the label "variable bias" in the figure. This is to say that the bias in the feedback is a factor of the extent to which the respondent understands her or his interest to be best served by giving frank and honest feedback. Over time, organizations can earn this understanding in their constituents and as they do bias goes down – hence the term variable bias.

One way we have helped our clients to recognize feedback bias – and work to mitigate it – is by conducting parallel data gathering by ourselves and by the client. The results of this experience provide the empirical basis for the figure below.



> ... the bias in the feedback is a factor of the extent to which the respondent understands her or his interest to be best served by giving frank and honest feedback....

Bi-modal data collection: light and heavy

Having decided on a strategy with respect to anonymous vs. non-anonymous and self-collected vs. independently collected, the next step is to envision the overall Constituent Voice system that best fits the organization and the steps to build that system, including which data collection methods to use when. There is no single model here. Introducing a Constituent Voice system requires adaptation in the organization's performance management system. This is never simple and is usually the main determinant of the pace of development of the Constituent Voice system.

Having emphasized the need for customization and the fact that Constituent Voice always involves organizational change, the following elements provide a robust framework for work.

Continuous light-touch feedback data collection

The ideal type end-state of a Constituent Voice system is one in which the organization is getting feedback continuously through micro-surveys and other techniques at diverse touch points with its constituents, supplemented by other feedback collection that is not touch point based. An example of a non-touch point based data collection exercise might be a campaign inviting constituents to call a toll free number to trigger a short survey or signal their willingness to be interviewed by phone.

A wide range of data collection mechanisms can be used, but where surveys are employed they should not contain more than three to five questions. The organization may have 15-25 questions in its full list, so these are rotated systematically through the data collection process to ensure an adequate sample of responses to each question.

Where the data collection mechanism and the capabilities and predilections of the respondents make it possible (for example, when respondents are willing and able to type out answers), closed Likert questions are combined with an open question that invites the respondent to explain the reason for the answer, or even to suggest solutions. Responses to open questions are analyzed and categorized.

Given the low cost and ubiquity of cell phones now, text message (SMS), USSD (the technology that drives pay-as-you-go), and interactive voice response (IVR) surveys are becoming a serious option for many of our clients.

> “... there is considerable value to be gleaned from periodic, high-quality, in-depth research....”

Occasional in-depth research

Even where micro-investigations are in place, there is considerable value to be gleaned from periodic, high-quality, in-depth research. Many of our clients from time to time (typically every 18-24 months) supplement continuous micro-surveys with longer surveys administered by an independent researcher using rigorous techniques. We have found that these longer surveys can help to answer questions emerging from the micro-surveys, as well as address strategic priorities of the organization and set management targets for the future.

The comprehensive surveys include both quantifiable rating questions (using Likert scales) and open-ended questions where respondents can amplify on their scores. Where the surveys are anonymous, selected demographic questions enable disaggregation and comparison of responses.

There are additional important forms of more in-depth research. These include formative and summative evaluation studies – conducted independently or internally. Other possibilities are more elaborate observational studies, and participatory research initiatives. Many organizations can leverage the considerable social science resources of local universities, often at no or low cost.

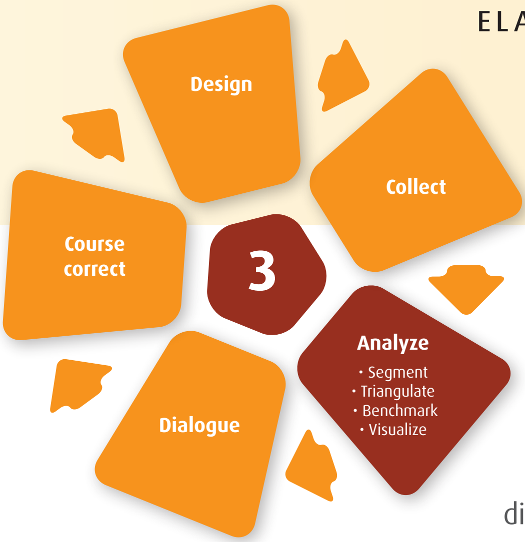
Data collection and research is not enough

The findings from continuous light-touch surveys and occasional in-depth research do not normally surface and develop solutions. They have powerful diagnostic value but in themselves are often not enough to guide a full management response. They signal where there are issues to address, and often provide clues to action. The micro-surveys will tell you if corrective measures, once taken, are working.

The development of effective solutions involves the other steps in the Constituent Voice cycle as described below under "Analyze" and "Dialogue". In "Analyze," we discuss how to compare Constituent Voice data with other evidence available, such as from normal project monitoring, or from other public data sets. In "Dialogue" we show how an organization's frontline workers can become the most valuable source of surfacing effective solutions.

> "... the micro-surveys will tell you if corrective measures, once taken, are working...."

Stage 3 • Analyze



So, we have asked the right questions, and collected valid answers, now we come to analyze and interpret the data.

Part of interpretation comes through dialogue with constituents to create a shared understanding of what the data and analysis may really mean. This "sense-making" aspect of interpretation is discussed in the next section.

Quantified perceptual data can be analysed using standard statistical methods to give actionable insight into the perceptions of different groups of constituents, and that by tracking these measures of perceptions and their analyses over time it is possible to refine ever more powerful insight. The emphasis in our analytics is to generate clear conclusions and to represent those findings in simple graphics. In addition to descriptive statistics, we employ three main types of analysis – segmentation, triangulation and benchmarking. We discuss each of these below.

Segment

Constituent Voice analysis typically begins by segmenting responses in two ways: by individual characteristics and by creating sub-groups of constituents (or clusters) depending on their responses. The general purpose here is to be able to use the analysis in order to develop corrective actions and strategies that fit with different segments of an organization's constituents.

Individual characteristics

To start, the client sets the importance to it for the individual (or, in the trade, demographic) characteristics of the respondents, for example, sex, age, nature of relationship with the client (occasional, frequent), location. By disaggregating answers to questions by demographic characteristics one can clearly see relative success and failure in connecting across the entire group of constituents.

Addressing the gender gaps

In Keystone's Development Partnerships Survey we are finding out that, even after considerable efforts by our clients to apply a gender approach, female respondents tend to report that they receive less capacity building support than male respondents. This points out that many INGOs need to redouble their efforts to consult with women and ensure their participation in capacity building support initiatives. We have recommended that clients strengthen monitoring metrics for women's participation in capacity building and going forward we will compare this project monitoring data with women's feedback.

Clustering

The other way is by clustering respondents according to their responses to certain questions in order to create categories that are more useful for developing and executing corrective actions. When we use the familiar 0 to 10 Likert scale in customer service surveys, we also use an approach to clustering called net promoter analysis (NPA).[5] NPA is disarmingly simple, but it has proven to be a reliable measure of customer loyalty, and a powerful lever for positive organizational motivation and change. As illustrated in the figure below, NPA classifies respondents into promoters, passives and detractors and calculates a single net promoter score (NP Score).

It may seem to some people that the division of respondents in these three groups is somewhat arbitrary. If we have for example a scale of "0- absolutely don't agree" to "10- absolutely agree", then why do we consider that someone giving a rating of 6 is a detractor? What the customer service industry has learned over time and what we are seeing in the social value creating sector, is that this accommodates the 'courtesy bias' in survey responses. Empirical evidence shows that people giving ratings in the middle or just above the middle of the scale are normally understating their dissatisfaction. The converse is not shown to be true. Those giving lower scores are not understating their satisfaction. In any case, if your goal is excellence, then you can never be satisfied with anything below a 9.

5 'Net Promoter' is a registered trademark of Fred Reichheld, Bain & Company and Satmetrix. For more see: **www.netpromotersystem.com**, as well as the open source net promoter community at **www.netpromoter.com**.

All Likert scales also lend themselves to clustering. The task is to combine scores in a way that creates groups that for the purpose of corrective action are likely to react similarly. For certain kinds of surveys (e.g., USSD or IVR), smaller scales are more appropriate and in some cases scale descriptors can allow respondents to identify themselves with a particular attitude, such as *doubter*, *fence-sitter*, *qualified supporter*, *unqualified supporter*.

One important sub-group are those who do not respond to surveys. Constituent Voice places a high value on response rates. Higher participation rates, while not always guaranteeing representation, reduce the likelihood of a significant response bias in the data. We track and report response rates as a key indicator of relationship quality. The companies that invest most seriously in customer satisfaction metrics do this. Some "customer-centric" companies even put non-responders into the detractor category when calculating net promoter scores. This pushes NP scores down, forcing them to work on winning high participation in feedback loops. This is discussed further under "Dialogue".

Because Constituent Voice is an ongoing process of engagement, over time we expect to see higher and higher response rates as constituents realize that great things happen when they provide frank and honest feedback. In order to understand more about sub-clusters, we routinely compare answers of two or more questions to determine the patterns of distribution of particular answers. So, we would ask, how do responses to a fairness question (*[...] treats me fairly*) compare with responses to an outcomes question (I *have more income and my family lives a better life because of [...]*). Applying standard statistical tests, we assign a statistical significance of any variations from what we would expect to see. For example, we might find no significant correlations for women as distinct from men with respect to either the fairness or outcomes questions when looked at independently. But when we do a gender disaggregation for how all detractors to the fairness question respond to the outcomes question, we might see (in fact have seen) that women report fewer benefits. These kinds of findings generate questions that are explored further in "Dialogue" activities with respondents.

The Meaning of satisfaction

A survey with tree plantation workers showed an unexpected pattern. Workers gave high scores on "satisfaction with the company" and low scores on "the company treats me fairly". When asked to explain these seemingly contradictory answers, the workers made it clear that their answers to the satisfaction question indicated that they needed the job, as they mostly had no alternative for wage employment. But, they added, this had nothing to do with how the company treated them, which they felt was too often unfair. By further enquiry we were able to surface specific grievances that, once addressed, resulted in a significant drop in arson in the company's forests.

> ... it is identifying patterns of responses that may show us previously hidden factors that are shaping results ....

We are not claiming that this kind of statistical analysis is generating definitive answers. Rather, it is identifying patterns of responses that may show us previously hidden factors that are shaping results. By teasing these out, we can use them in sense-making activities with constituents and more generally in the ongoing interactions between the organization and its constituents.

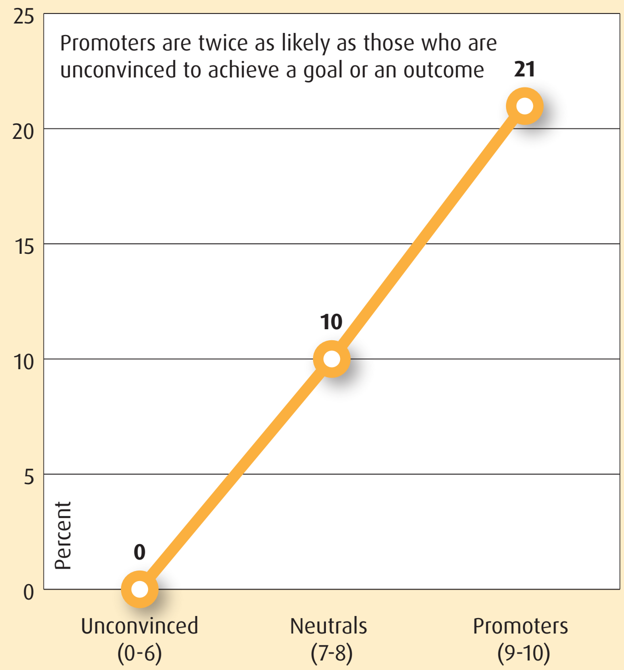
We want to strengthen relationships to get better results for constituents and for the organization, so the focus is always on testing actions to improve. Having discovered, for example, that 30 percent of those we serve are consistently detractors, we can we address their specific concerns. The data identify them for us, and they will tell us when we have succeeded with them. Data throw up questions or hypotheses; we turn these into a process of shared enquiry with constituents. The operative word here is “we”, and the data tell us much more than we knew before about the “we”.

Triangulate

Triangulation is a fancy word for comparative analysis, usually drawing from different data sources or using different methods of research (classically, quantitative vs. qualitative).

For example, in-depth interviews could be conducted with certain groups to gain deeper insight into their perspectives on program outcomes. These perspectives are then compared with the findings from survey questions on outcomes. Putting the two together underwrites higher reliability of the findings.

A primary objective in much of our client work is to discover consistent correlations between real time perceptual data from our surveys and later occurring outcomes measured independently (and preferably objectively). The customer service industry provides the most famous example of this. After fifty years and innumerable studies, no one doubts that customer loyalty properly measured in surveys today is an accurate predictor of profits, shareholder value and corporate growth. Organizations who triangulate perceptual feedback data with outcome data are beginning to find some interesting correlations. One Keystone client who has begun comparing the two sources of data have found that promoters are twice as likely to achieve an intended outcome



than those giving low scores. (see figure above). By asking a handful of very short questions regularly, the organisation is able to identify today those of its constituents who absent some change are not going to achieve intended outcomes in the future. Based on these metrics, the organisation is providing extra support to these individuals.

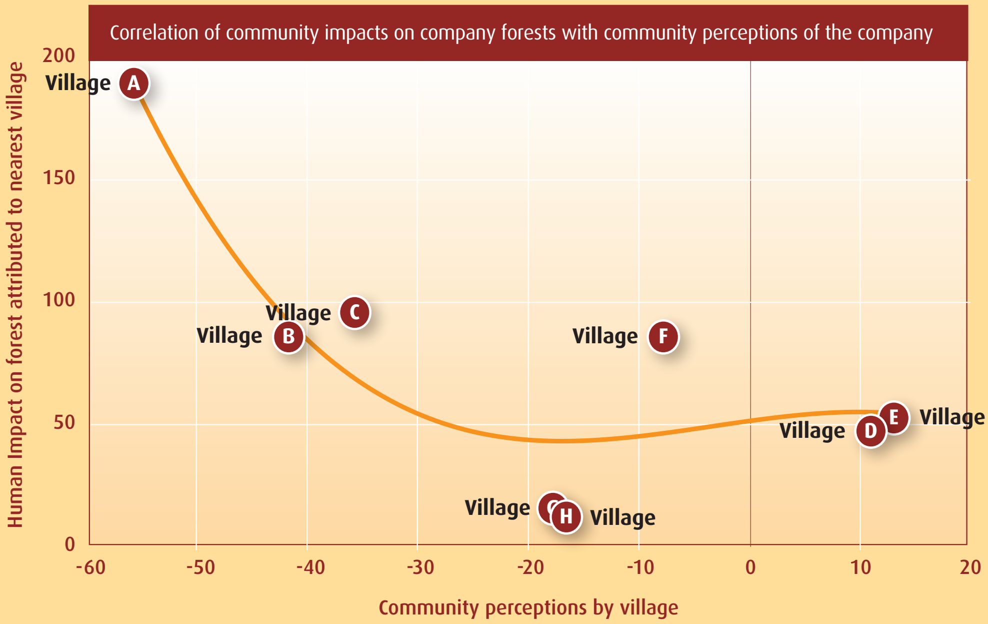
Triangulation is proving to be a powerful tool, helping anticipate outcomes as well as enabling understanding on how to improve programs and interventions to maximize impact.

We are also seeing some exciting predictive indicators emerging in other areas of our work. Student feedback correlates to the scores they later receive on standardized tests. Worker feedback correlates to absenteeism, accidents, grievance rates and incidents of sabotage. Smallholder farmer feedback anticipates rates of adoption of new technologies and subsequent crop yields. Responses to a question about connections to community predict how quickly a very low income person progresses in attaining financial goals. Feedback from displaced people living in camps points to the obstacles to their self-reliance and how to remove them. Predictive indicators are one of the best tools available to achieve better results more quickly, particularly when combined with benchmarking (as discussed below).

In addition, feedback from constituents can be compared to other metrics such as financial performance, census data or data from other sources. In a youth employment project for example we may cross feedback from youths expressing that they perceive

Work by day, burn by night

In the same case of the tree plantation company, the company gained confidence in the findings from the survey when we compared the survey findings disaggregated by village with the company's own records of harm to its forests. The table below shows that there was little harm to the forest around the villages that reported more positive attitudes to the company, while those that reported more negative scores were located next to areas of significant harm to the forest.



themselves to be more employable with data from the local employment office indicating an increase in applications to social insurance for people in the same age range.

Benchmark

Quantification of qualitative feedback data enables us to create performance benchmarks – not on the basis of technical 'ratings' by external inspectors, but on feedback from those in the best position to speak of their experience of an organization or service. Comparable feedback data are generated when the same questions are asked in the same ways across groups or at different times. Some organizations compare their own performance over time or across units but only rarely do they do so against other organizations doing similar things.

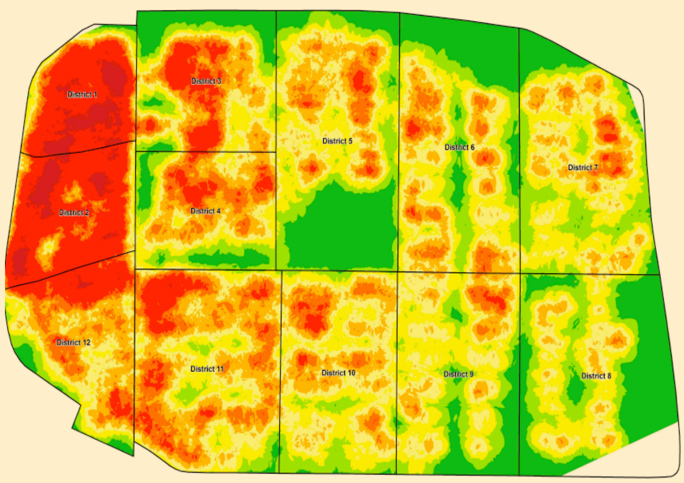
Benchmarking deepens your understanding of what a particular numerical answer means by showing it in relation to other scores. By comparing against the average, you understand what normal means. And by comparing against high-performing outliers, your imagination of what is possible is stretched. Say the manager in Region A sees that his region received a net promoter score of 25. He might be content with this until he sees that Regions B-G have an average NP score of 37. A close look at Regions B-G also reveals that the top scoring Region F has an NP score of 65, including only 2 percent detractors. The manager in Region A now knows what top performers can do to lift scores, and can ask the manager in Region F how she did it.

In addition, when a critical mass of organizations start to pool their feedback data and benchmark against each other, we can identify trends in the sector as well as extract best practices that can be used for improving performance collectively.

Keystone offers a data sharing opportunity to clients to enable them to compare their data and benefit from field level insight deriving from their aggregated feedback data. In partnership with leading grantmaking foundations and Feedback Labs, Keystone has created the Feedback Commons as an online platform where organizations can "share and compare" their feedback data.

Visualize

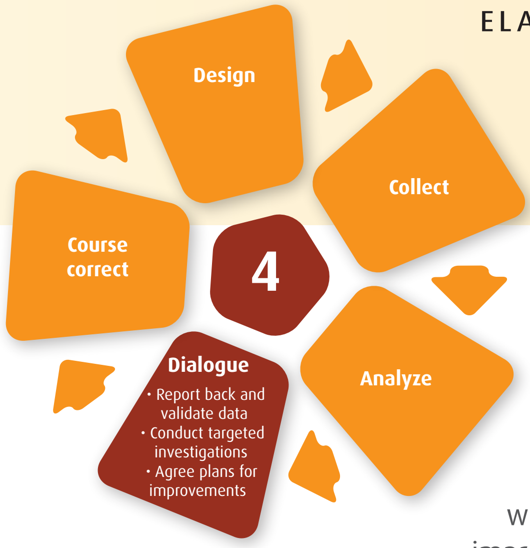
There are a growing selection of sophisticated and easily available tools that can help visually present data. Visualizations can help deepen understanding of the data, and identify meaning within it. For example, geo-mapping tools plot data across a geographical area. The example visualization below plots scores within a geographical area of a refugee camp. It uncovers a concentration of low scores (shown in red) in a particular area that is hard to ignore! Such tools are particularly useful in dialogue sessions with respondents.



Towards indexes

Net promote analysis is one way to reduce complex patterns to simple, yet powerful single metrics. Another technique is indexing. We have found that at times it can be useful to combine responses to questions into a single index that is tracked and compared over time. The value of an index is in its ability to represent separate but related elements in a single number. Reducing data into bite-size pieces helps organizations understand it and respond to it. The more we know about how individual questions correlate to performance and results (as discussed above under "Segmentation"), the more explanatory the index.

We have found it particularly useful to aggregate very specific micro-questions relating to service quality at a specific touch point into a Service Quality Index, for example. When different services offered by an organization are indexed in this way, it is meaningful to compare aggregate service quality across these services. Similarly, by combining Relationship Quality questions we achieve a Relationship Index that clients have found to be a compelling motivator for staff to implement reforms.



Stage 4 • Dialogue

"It's not about the data, stupid."

The single most important step in the cycle involves closing the loop by reporting back to constituents what you heard and co-creating solutions. We call this activity sense-making, as in making sense of the findings through open conversations with respondents that generate a shared understanding of ways to improve.

Two things happen when you ask constituents questions. One is you get answers. The other is that you raise expectations. By holding a dialogue with respondents you do two things. You refine and validate the meaning of the answers to the questions. And you manage the expectations created by asking the questions by emerging solutions that are within the actual capabilities of the organization and the respondents.

In the place of unbounded expectations, you now have a two-way dialogic process that grounds constituent expectations in granular data – "This is what we heard you saying. This is what we propose to do. What do you think of that? And by the way, these are some of the factors that constrain our capacity to respond here. What can you contribute to a solution here?"

Organizations that move beyond data collection to dialogue not only learn and improve their performance, they underwrite higher response rates and more frank feedback in future surveys because constituents clearly see the value in providing feedback.

In fact, this Dialogue stage is crucial for investing in the feedback mechanism. By clearly demonstrating that feedback is being listened to and responded to appropriately, an organization will yield higher responses and even better feedback data as it continues to repeat the cycle. The next stage – Course correct – will provide the opportunity to tweak the feedback mechanism based on the discussions that happen during this Dialogue stage.[6]

6 An example from a Ugandan program of community monitoring of rural health clinics illustrates this general point in a fascinating way. An experimental study has found that "[e]fforts to stimulate beneficiary control, coupled with the provision of report cards on staff performance, resulted in significant improvements in health care delivery and health outcomes in both the short and the longer run. Efforts to stimulate beneficiary control without providing information on performance had no impact on quality of care or health outcomes." (**http://econ.worldbank.org/external/default/main?pagePK=64165259&theSitePK=469382&piPK=64165421&menuPK=64166093&entityID=000158349_20140826090814&cid=DEC_PolicyResearchEN_D_INT**; downloaded 5 September 2014). In other words, when constituents were given information about organizational performance, the whole community monitoring effort was more effective in generating improvements in outcomes. This is an original and illustrative insight into how to invest in the feedback mechanism.

> …Two things happen when you ask constituents questions. One is you get answers. The other is that you raise expectations….

Clear data visualizations – like the geo-mapping example earlier – are powerful aids to dialogue sessions. Net promoter analysis and benchmarked data are also easy to visualize in ways that constituents can quickly sink their teeth into. They are accessible and intuitively powerful. They ground conversations in empirical descriptions that could otherwise be hijacked by the loudest voices in the room. At another level, by holding oneself to account by being transparent about feedback received, one builds confidence and trust and enhances the credibility of programs.

Broadcasting and publishing

There are many different ways of reporting back and discussing feedback. The mix of methods employed depends on your objectives. One-way broadcast and publishing modes are often used as the first step to get the main findings out there and establish a basis of transparency and accountability. They set the stage for more probing efforts.

There are many well-known and simple broadcast methods and the context will usually quickly determine what is most cost effective. Some of the techniques that clients have used effectively include: distributing printed reports, screening video reports, community radio, call-in recorded messages, posters, and interactive screens on the walls of waiting rooms.

Group discussions and key informant interviews

Meetings or focus groups can be structured as reporting back on survey findings and open enquiries into the questions or hypotheses arising from the data. Some organizations prefer independent facilitation at such report back sessions. Others ask staff to add Constituent Voice dialogues into their normal activities.

In many dialogues we have made good use of the readiness question – *"I feel that am ready and willing to try new things offered by […]"*. All relationships are two-way affairs and by asking this question we are encouraging constituents to look at what they are bringing to the relationship with the organization. When we compare survey answers to the readiness question to respondent ratings of the organization's relationship quality we usually see that respondents rate themselves higher than they rate the organization. We have found that putting this difference on the table at sense-making meetings generates searching and useful conversations.

Interviews – by phone or in person – with 'key informants' offer another way to explore a set of questions emerging from data collection. Key informants are people who have been identified as being willing and able to explore behind their answers in surveys. We often invite people to volunteer for such follow-up interviews in the survey stage. When used alongside focus groups, key informant interviews can also be used to crosscheck what is emerging in the focus groups.

Informal follow up investigations

The simplest way to discover the answers to questions arising from the micro-surveys is often to ask front line staff to discuss these questions with constituents informally as part of their regular interactions. For example, micro-surveys may show that younger women are like other respondents in most questions, but are consistently less trusting of the organization. This begs the question why. To dig deeper, the organization may launch a time-limited micro-investigation in which front line staff share this finding informally with female and male constituents and keep a record of their possible explanations. These are collected and analyzed to see if a clear pattern emerges as well as a consensus about possible corrective actions.

Are women tougher graders?

A US anti-poverty organization found that women were giving the organization lower scores than men. Project monitoring data suggested, however, that women and men were benefitting equally and behaving similarly based on metrics like longevity in the programs. The organization raised the pattern in small focus group sessions with those they served and three hypotheses surfaced:

- Women may have more struggles (or example as single parents) than men and so need more support overall. Perhaps the organization can't cover everything. Support may need to come from a variety of places.
- Perhaps women come in with more "tools" in their toolbox and therefore the organisation contributes less to this. For example, perhaps they already feel "able to stand on their feet and achieve what they want" (one of the questions with a strong difference between men and women scores) and so the organization does not contribute to this.
- Are men afraid/unwilling/too biased to give low scores to our mostly female volunteer-base?

One staff member called this last point "the flirtation factor". So it may be that it is not that women were tougher graders, but that men were grading to impress. The organization is now asking new questions in surveys to test these hypotheses. It is also adding staff gender to its data set to see if disaggregating by the sex of the staff providing the support validates the overall pattern of lower scores from women service recipients.

> … this approach creates value in two ways. Firstly, constituents appreciate the value placed by the organization on their feedback. Secondly, staff members are empowered to become actively curious, to be evidence-based problem solvers for the organizations …

In addition to generating possible solutions this approach creates value in two ways. Firstly, constituents appreciate the value placed by the organization on their feedback. Secondly, staff members are empowered to become actively curious, to be evidence-based problem solvers for the organizations. Employees who excel at this can be recognized and rewarded.

The most customer-centric companies excel at this kind of post-survey investigative activity. Some, for example, routinely call those who give a 6 or less to the net promoter question within 48 hours. In effect, they treat a low survey score as a formal grievance, and act to address it.

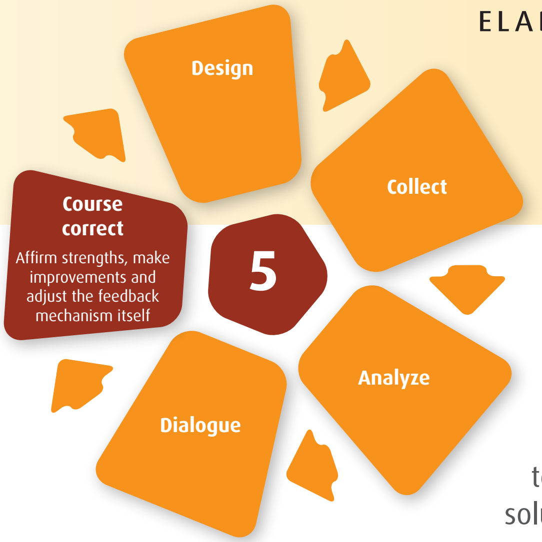
Co-creation

One of the key outputs of the Dialogue stage is a plan for course-correction, the next stage in the cycle. Such a plan is not only best when it comes after these further discussions with constituents, but also when it actively includes constituents. Using the feedback data as the starting point for an open dialogue on how best to improve will allow organizations to develop solutions with and alongside those are meant to benefit from them. As we know, development is not something that can be done to people, it relies on their involvement, and organizations often find the best ways to help constituents when the co-create solutions. Later cycles of the feedback system will confirm whether these were the right changes, or if there are further improvements that can be made.

Public reporting

The strong trend toward transparency and accountability in public reporting offers practitioners of Constituent Voice a unique opportunity to leap frog to the next generation of accountability reporting, which will see annual reports replaced by real-time performance data flows, including Constituent Voice data. Organizations publishing in real-time will lead the next wave of sector-wide dialogues on performance standards and benchmarks.

Stage 5 • Course Correct



Use it, or lose it!

Constituent Voice is an iterative performance management tool, not an exhaustive research study. It tells you when you have an issue. It empowers managers and especially frontline staff to investigate, experiment and engage constituents to find lasting solutions. And it will tell you whether you have solved the issue.

Development of any kind is a process requiring constant change, and Constituent Voice gives you the ability to get back on track when something has gone awry. All organizations get it wrong at times, and you'll fix those mistakes by honestly taking stock throughout the Constituent Voice cycle and using the data to change the way you do things.

Armed with empirically valid data, Constituent Voice enables organizations to secure buy-in across staff and those you serve and create a culture of small, iterative steps to test your way forward – a culture based on improvement rather than blame. This culture based on accountability and transparency is the ultimate guarantor of performance and results over time.

The other purpose of the course correct step is to tweak the whole Constituent Voice system itself. This is a cyclical return to stage 1: Design. Use feedback from respondents to rethink your data collection tools and methods for analysing the data and how you engaged with constituents during stage 4: Dialogue. This is the chance to improve the system and think about how to further maximise the value you get from the process.

Some concluding observations

The main objective and value of Constituent Voice is to create and sustain an organizational culture of continuous improvement grounded in authentic conversations about performance with those you serve. But to round off this technical note, we want to highlight the flexibility of Constituent Voice systems by sharing some of the supplemental ways that Constituent Voice pioneers are using it.

Testing your theory of change or new proposals

Once an organization has a Constituent Voice system in place, it can be used to test hypotheses in the theory of change, or to explore other important questions that come up. For example, an education provider may have a hypothesis that parents who have more voice will be more likely to volunteer in specified activities. A simple experiment that tracks voice scores to patterns of response to volunteer opportunities provides a compelling answer. By eliminating other explanations through other Constituent Voice scores and perhaps through other investigations, it is possible overtime to pinpoint the key causal mechanisms for parental volunteering.

Another way to re-purpose Constituent Voice that is increasingly common is in planning. Constituent Voice data collection can be used to poll opinions on different proposals. This can be as simple as the best time to hold a meeting to a fundamental change in organization strategy.

The value of self-efficacy

With its basic Constituent Voice system in place and yielding consistently helpful metrics, one client has started to ask questions in micro-surveys that will be part of its larger theory of change. This organization hypothesizes that if it can strengthen its clients' self-efficacy – belief in one's own ability to complete tasks and reach goals – that they will in fact become better at completing tasks and reaching their goals. By adding a question to its CV system that tracks self-efficacy on a 0-10 scale (To what extent do you agree with the statement, "I believe I can achieve my goals"?), and revisiting that question over time, it will be able to compare time-series answers for individuals against their actual goal progress as recorded in the organization's normal client data monitoring. In this case, while it is still early days, the first few months of data collection are strongly affirming the self-efficacy hypothesis.

As part of your evaluation framework

Constituent Voice data contributes to wider evaluation. Both summative and formative evaluations are powerfully enabled by Constituent Voice data. Evaluation costs are reduced to the extent that evaluations can rely on Constituent Voice data rather than on collecting new data. Constituent Voice systems have the intrinsic effect of strengthening normal activity monitoring because their emphasis on short cycle iteration makes regular calls for comparison with monitoring data.

Communicating across constituents

The information and insight generated through Constituent Voice are often of great interest across constituents. Constituent Voice systems allow organizations to be a nerve center for communication of perspectives across the ecosystem around an issue or program.

Feedback from service recipients, for example, is of interest to the wider field of peer organizations and of course to investors and donors and government. Similarly, service recipients can make good use of the feedback from other constituents. Parents of schoolchildren have views on the school that the Ministry of Education wants to hear. And the Ministry's policies and positions can often touch directly on the powerful interests of parents. Since parents often have no idea what theses policies and positions are, Constituent Voice provides a way of finding out that is directly linked to ongoing closed feedback loops.