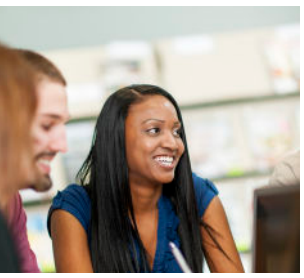
FUND FOR SHARED INSIGHT



LISTEN FOR GOOD

2016-2017 COHORT FINDINGS

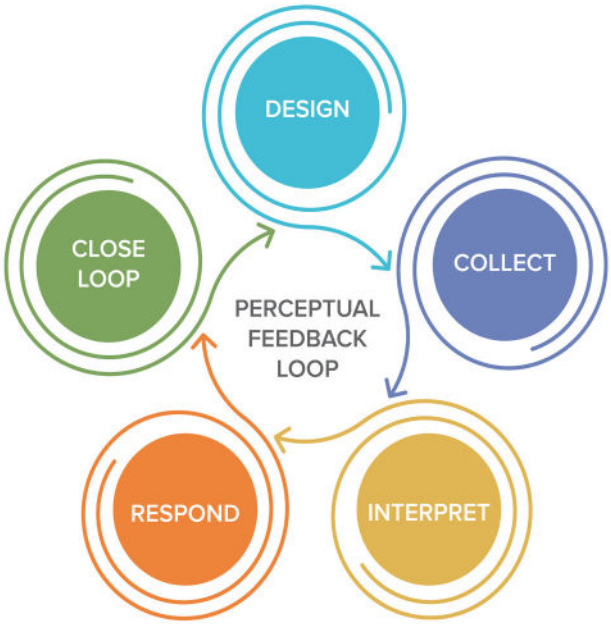
ORSIMPACT

CC BY NC SA

Introduction

Listen for Good (L4G) is an initiative of Fund for Shared Insight (Shared Insight) dedicated to building the practice of listening to feedback from the people nonprofits and foundations seek to help. L4G supports nonprofit partners in implementing a five-step feedback process designed to ensure high-quality feedback loops that lead to meaningful change, which includes designing the survey; collecting, interpreting, and responding to the data; and closing the loop with clients (Figure 1). To gather this feedback, L4G nonprofit partners use a semi-standard survey instrument, including the Net Promoter System NPS®, as a simple, systematic way of listening to the people at the heart of their work.

Figure 1 | Steps in perceptual feedback loops



Since 2016, Shared Insight has made a total of 158 L4G grants to nonprofits, nominated by 69 co-funders over three grantmaking rounds. These nonprofits have administered more than 60,000 client surveys so far. The grants are 2:1 matching grants over two years, where co-funders provide one third of the grant money, while Shared Insight provides the remaining two thirds. In addition to the grants, L4G support includes a subscription to SurveyMonkey, a set of semi-standard survey questions, access to sector-specific performance benchmarks, and technical assistance (TA) through 1:1 coaching, group learning, and a resource website.

The initial grantmaking round started with a cohort of 46 nonprofits in 2016, with the subsequent two cohorts beginning the process in September 2017 and January 2018. As L4G has evolved, there have been changes to the model, creating notable differences between the three cohorts. While the 2016 cohort received $60,000 grants, the 2017 and 2018 cohorts received $45,000. In addition, L4G supported the 2016 cohort through a team of two TA providers, and expanded to a greater number of TA providers in 2017 to better serve a growing number of organizations (66 instead of 46). Finally, the 2017 cohort received earlier and more frequent access to webinars and support materials.

Introduction

In addition to supporting nonprofits, L4G has sought to experiment and learn about applying NPS® in the beneficiary context; accelerate the building of infrastructure for feedback practice; engage funders in supporting feedback practice; and capture and share lessons learned to positively catalyze the feedback field. With the goal of supporting a greater number of organizations in implementing high-quality feedback practice, L4G will launch three test rounds between late 2018 and 2019 where new nonprofit partners will engage with a web-based support system to implement the five-step feedback loop. The service will ultimately open to the public in 2020 with the goal of supporting a greater number of organizations in implementing high-quality feedback practice.

Over the past three years, ORS Impact has served as learning and evaluation partner to Shared Insight and L4G. As part of this work, ORS Impact has collected, analyzed, and reported data from L4G's different nonprofit partner cohorts. This report provides an overview of our assessment for how L4G has performed as a whole during these first three years. By analyzing data collected from 104 out of the 112 nonprofit partners from the 2016 and 2017 cohorts[1], we reflect on questions such as:

- How is the L4G model working for nonprofit partners?
- What impact is L4G having on organizations?
- How does L4G relate to organizations' equity, diversity, and inclusion work?
- What type of organization finds the most value in L4G?
- What implications do the findings have for L4G and Shared Insight?

In exploring and answering these questions, we provide a higher level view of how L4G has worked across the two grantmaking periods, and identify key lessons and implications for new L4G partners under the current model, for the test rounds and public-facing version of L4G, and for evaluation efforts moving forward.

Data Sources & Methodology

For this report, we analyzed online survey data collected from nonprofit partners within the 2016 and 2017 cohorts, as well as conducted phone interviews with representatives from 16 co-funders who nominated the 2016 nonprofits. These surveys and interviews sought to explore organizations' perspectives on their experience with L4G, the changes they have made or experienced since the beginning of the L4G grant, and their learnings about implementing high-quality feedback loops or funding feedback practice.

To analyze quantitative data, we considered descriptive statistics and compared responses from different types of organizations through statistical significance tests. In these tests, we analyzed data for statistically significant differences between cohorts (2016 vs. 2017), timing of survey (six, 12, or 24 months after participating in L4G), organizational budget size, area of service, and nature of interaction with clients. Figure 2 provides an overview of different categories of organizations used for the analyses. We identify statistically significant findings throughout the report as "significant" or "significantly" different responses. For open-ended questions, we used a thematic analysis approach to code the data for prevalent themes across responses. Appendix 1 provides greater detail about data sources and analysis methodology.

[1] We have not collected data from the 2018 cohort yet, as they are too early in the process.

Figure 2 | Overview of categories of organizations used for the analyses

Budget size
< $1 million
$1-$5 million
$5-$10 million
> $10 million

Nature of client interaction	
Indirect	Organizations that interact with clients "behind the scenes," through intermediaries.
One-off	Organizations that interact with clients only once or, if recurrent, tend to be transactional in nature, with limited opportunity for relationship-building with staff.
Sustained	Organizations that interact with clients over a longer period of time in which relationship-building is a primary goal of the interaction.

Area of service
Health
Education
Community and Economic Development
Human Services
Public Safety
Arts and Culture
Environment
Public Affairs

2016 Cohort Surveys

The 46 nonprofits from the 2016 cohort have now had two years to implement feedback practice through L4G. We surveyed these organizations at three time points: six months, 12 months, and 24 months into their grant period. At each time point, we collected data from both agency leaders and program managers. As a result, we have a total of six surveys providing insights about the 2016 cohort's experience with L4G.

2017 Cohort Surveys

The 66 nonprofits in the 2017 cohort have been implementing feedback loops for nine months. We have collected data from these organizations once so far – six months into L4G – and surveyed only program managers. Figure 3 shows response rates from 2016 and 2017 cohort surveys.

Figure 3 | Response rates from 2016 and 2017 Cohort surveys

	2016			2017
	6 months	12 months	24 months	6 months
Program Managers	78%	87%	83%	88%
Agency Leaders	67%	85%	72%	N/A

L4G's Impact on Nonprofits

This section examines how and to what extent L4G is having an impact on nonprofit partners, and if there are certain types of organizations that find more impact from L4G. In addition, we reflect on how L4G relates to organizations' equity, diversity, and inclusion work.

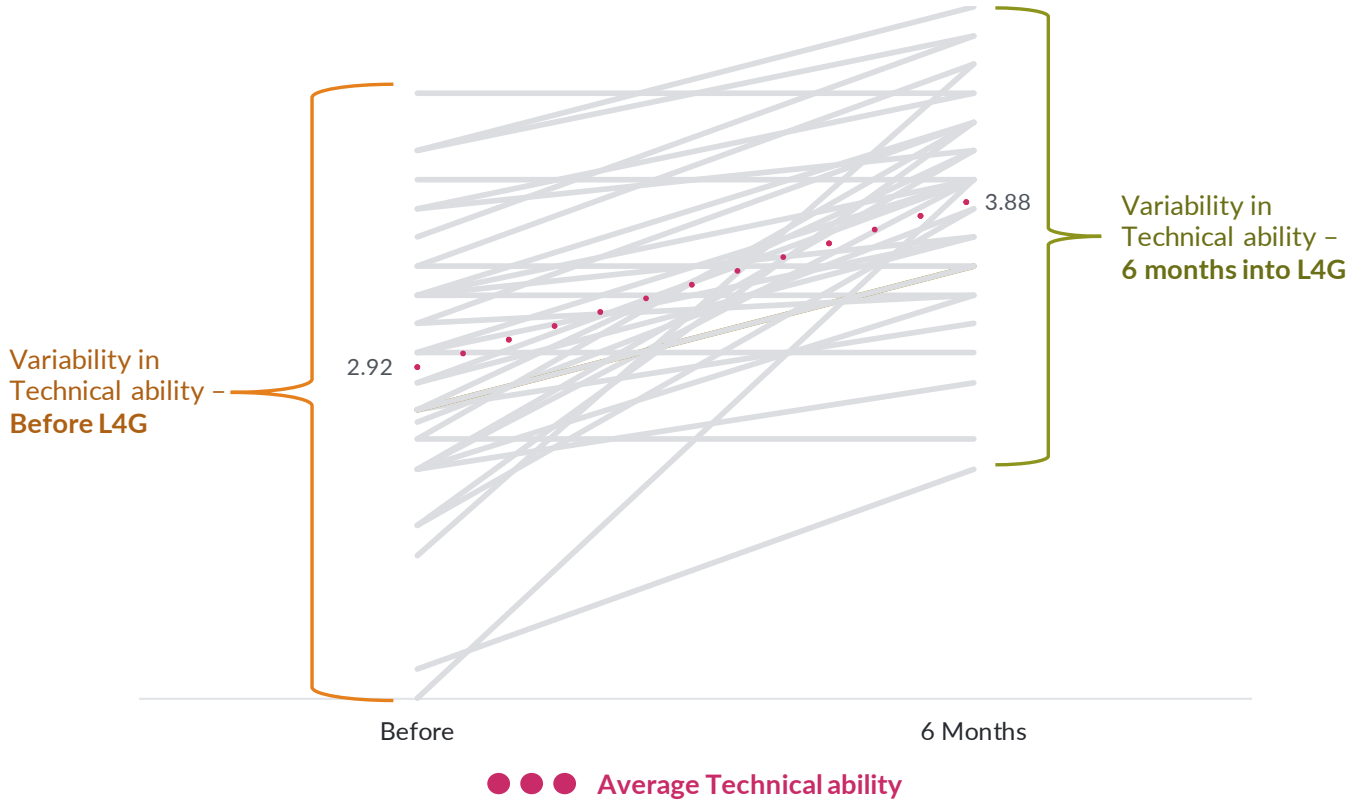
Organizations from both the cohorts report improvements in their ability to implement feedback-related tasks, but there may be a ceiling effect to this growth.

Throughout their L4G grant, both program managers and agency leaders have reported improvements in their organization's technical ability to carry out six different tasks associated with high-quality feedback loops, including: implementing surveys with clients at least twice a year, achieving high response rates, collecting useful data from clients, analyzing the data, interpreting the data to inform their work, and closing the loop with clients.

The data shows that before participating in L4G, there is greater variability in the levels of ability to implement these tasks across organizations; however, after six months of participating in L4G, the variability across organizations decreases. This suggests that L4G may be creating a level playing field of technical ability among organizations. This pattern is consistent across time points with variability decreasing from before the L4G grant to six months after, six to 12 months after, and 12 to 24 months after.[1] This overall change in variability across time is statistically significant for the 2016 cohort (Figure 4).[2]

[1] Ability is measured as self-reported ability, and organizations report their perceived ability from before L4G engagement at the 6 month mark, along with their perceived ability 6 months into L4G.
[2] The variability is significantly different across time points using the average scores between program managers and agency leaders as a measure of ability for each organization ($p < .05$).

Figure 4 | Variability in levels of ability to implement tasks before L4G and 6 months after L4G, 2016 cohort[3]



In addition, organizations from both cohorts report significant improvements across all tasks from before the L4G grant to the six-month mark.[4] At 12 months, program managers from the 2016 cohort continue to report significant improvements in their ability to perform five out of the six tasks (all except achieving high response rates).[5] Meanwhile, also at 12 months, agency leaders report significant improvement in two out of the six tasks (implementing surveys at least twice per year and analyzing data). However, when moving from 12 to 24 months, neither program managers nor agency leaders report significant improvement in any task, and both groups identify a significant decrease in one task: the ability to implement surveys twice a year.[6] This pattern suggests that there may be a ceiling effect between six and 12 months, whereby organizations learn a lot up to that point, but their self-reported ability does not continue to grow at the same pace afterwards (Figure 5).

[3] N = 46. Each line represents program manager data from each organization.
[4] All p-values < .01.
[5] p-values range from < .01 to <.06. A p-value < .06 is considered to be approaching significance.
[6] Agency leader p-value < .05; program manager p-value < .05

Figure 5 | Change in ability to implement feedback tasks over time, 2016 cohort[7]

Organizations with higher ability to implement feedback loops are more likely to see impact on program effectiveness.

To better understand L4G's impact on organizations, we considered which factors could affect the extent to which organizations saw an impact on program effectiveness and found that technical ability is significantly and positively related to reported impact on program effectiveness.[8] In other words, as organizations report increases in their technical ability, they report that L4G has a significantly higher impact on programmatic effectiveness. This finding is encouraging since L4G is significantly improving organizations' ability to implement feedback loops. Given the importance of the increase in organizations' ability, we analyzed whether an organization's characteristics were related to its ability. The additional analysis shows that organizations with sustained interactions reported significantly higher ability scores than organizations with one-off interactions in their 24-month survey.[9]

[7] N = 46. Each line represents program manager data from each organization.

[8] This finding is based on 24-month program manager data in a regression model controlling for the effects of budget, nature of interaction, area of service, insights gained, and L4G effects on internal organizational factors (e.g., values, culture, etc.). After controlling for these effects, program manager ratings of ability uniquely predicted perception of L4G impact on program effectiveness ($p < .05$).

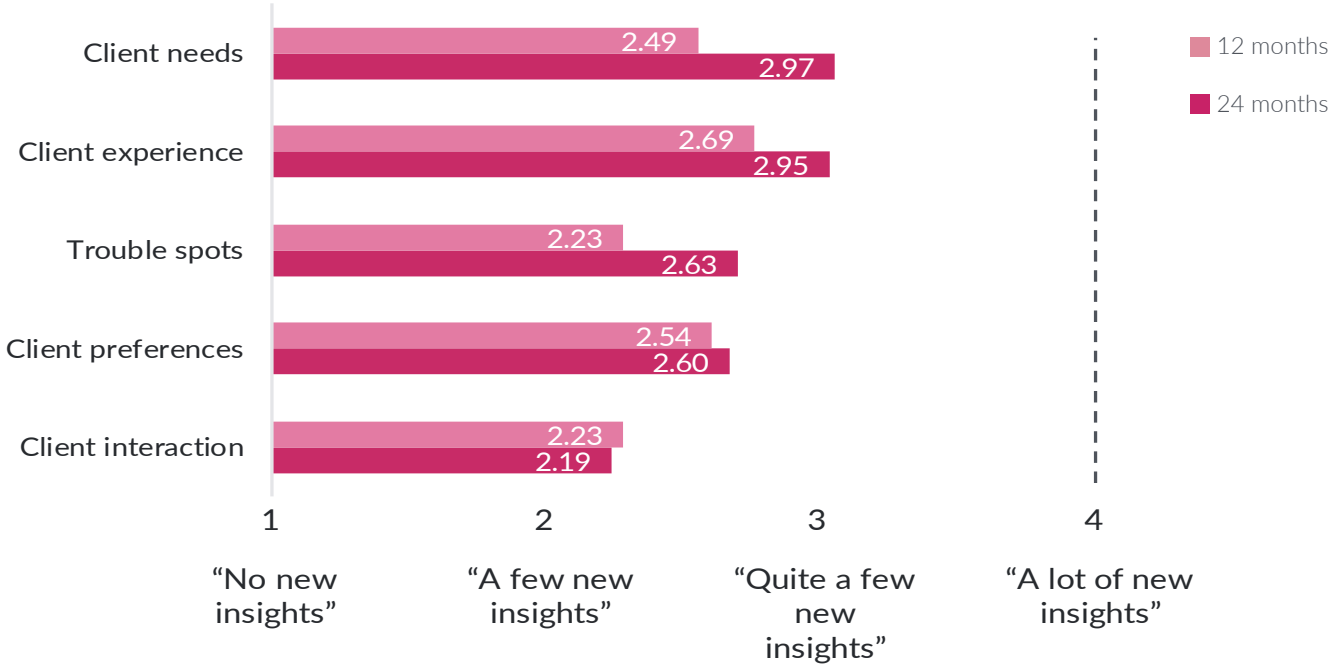
[9] $p < .05$, after controlling for the effects of budget, indirect interactions, and area of service on program manager rating of ability.

L4G provides organizational insights through both new information and data to confirm prior beliefs.

Organizations are gaining insights from the information they collect, and those insights serve two equally important purposes: to generate new knowledge and to confirm prior beliefs with data, providing a more rigorous and precise understanding of hypotheses held prior to conducting feedback practice.

When considering what organizations are learning through these insights, data from the 2016 cohort at 12 and 24 months suggests that organizations have gained the most insights about clients' needs, trouble spots, and reasons behind differences in client satisfaction. On the other hand, they have gained fewer insights into how clients interact with the organization. Another emergent pattern is that organizations report gaining more insights at 24 months than at 12 months. Although the increase is not statistically significant at this point, the pattern holds true in four out of five areas measured in both time points (Figure 6), and is noteworthy for analysis in future check-ins.

Figure 6 | New insights gained by organizations across different areas of their work.[10]



Survey data from the 2016 cohort at 24 months reveal that 46% of program managers are finding that the insights provide a balanced mix of new information and confirmations, 43% are mostly finding confirmations of prior beliefs, and 11% report finding mostly new information from the L4G process.

[10] 12-months N = 40; 24 months N = 38

Organizations that gain more insights are more likely to see impact on program effectiveness.

As we continued to analyze the factors that could affect the extent to which organizations see an impact on program effectiveness, we found that gaining insights is significantly and positively related to program effectiveness.[11] In other words, as organizations gain more insights from the data, their programmatic effectiveness significantly increases. Furthermore, when looking at relationships between gaining insights and other organizational characteristics, we found that organizations with smaller budgets and sustained client interactions are more likely to gain insights from L4G.[12] Specifically, smaller organizations – those with budgets of less than $1 million – reported gaining significantly more insights than organizations with larger budgets. Similarly, organizations with sustained client interactions also report gaining more insights than organizations with "one-off" interactions, or those who act as intermediary organizations without direct client interaction.[13]

Organizations are translating these insights into changes to better serve clients.

According to program managers, organizations have made and planned changes in different dimensions of their work, including program offering, operations, new services, and staff-client interactions. The proportion of organizations in the 2016 cohort reporting that they have made changes has increased between the 12 and 24-month surveys, most notably in program offering and operations (Figure 7). Not surprisingly, most organizations reported making changes in the programs where they collected feedback (78%); however, there were some ripple effects from the feedback, as 26% of organizations reported also making changes in other organizational activities, while 17% made changes in programs where they did not collect feedback.

At the 12-month mark, according to program manager responses, organizations with sustained client interactions were more likely to have made changes to program offering than organizations with one-off or indirect interactions.[14] This pattern remains true at the 24-month mark, this time in agency leader responses.[15] While these findings are statistically significant at both 12 and 24 months, the small number of organizations with indirect services may be a source of error in the statistical test. Nonetheless, the pattern remains relevant and noteworthy to analyze in future surveys.

---

[11] This finding is based on 24-month program manager data in a regression model controlling for the effects of budget, nature of interaction, area of service, ratings of ability, and L4G effects on internal organizational factors (e.g., values, culture, etc.). After controlling for these effects, program manager insights gained uniquely predicted perception of L4G impact on program effectiveness ($p < .1$). This is considered to be trending toward significance.

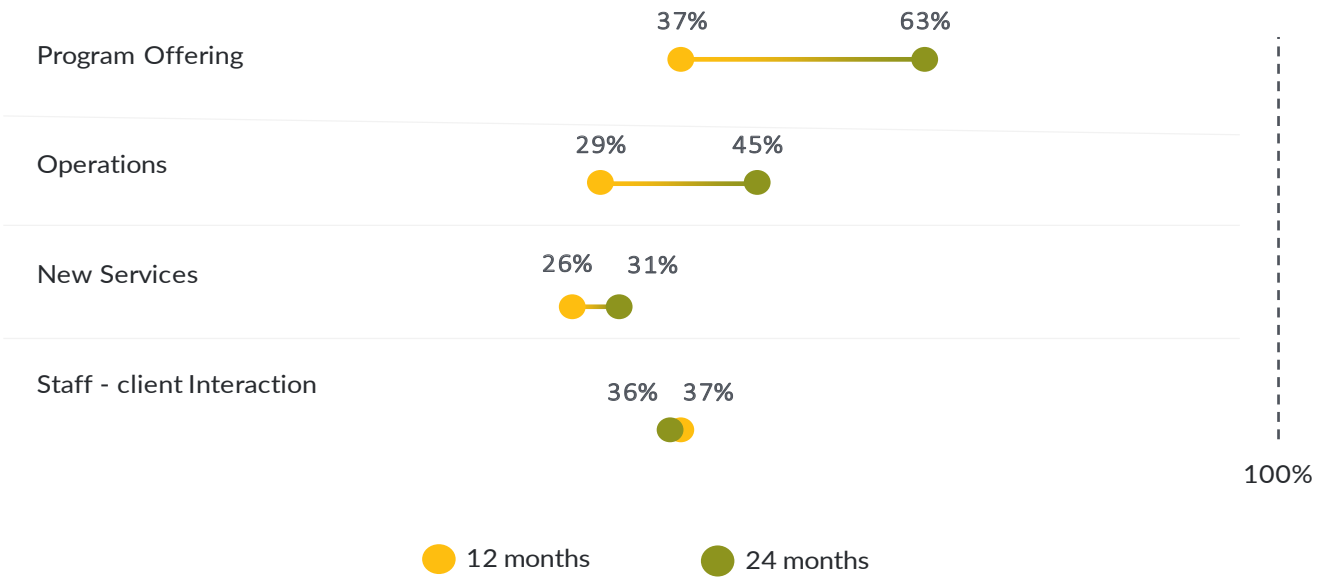
[12] $p < .05$ for organizations between $1M and $5M, $p < .05$ for organizations between $5M and $10M.

[13] $p < .01$

[14] $p < .01$

[15] $p < .05$; program managers were chosen to answer this question at 12 months, while agency leaders were chosen at 24 months.

Figure 7 | Percent of organizations making changes as a result of L4G insights at 12 and 24 months.[16]



[16] 12-month survey N = 35; 24-month survey N = 40

L4G and Equity, Diversity, and Inclusion

L4G advances organizations' equity, diversity, and inclusion work by providing clients a seat at the table at various steps in the process and allowing them to shape the services that benefit their communities.

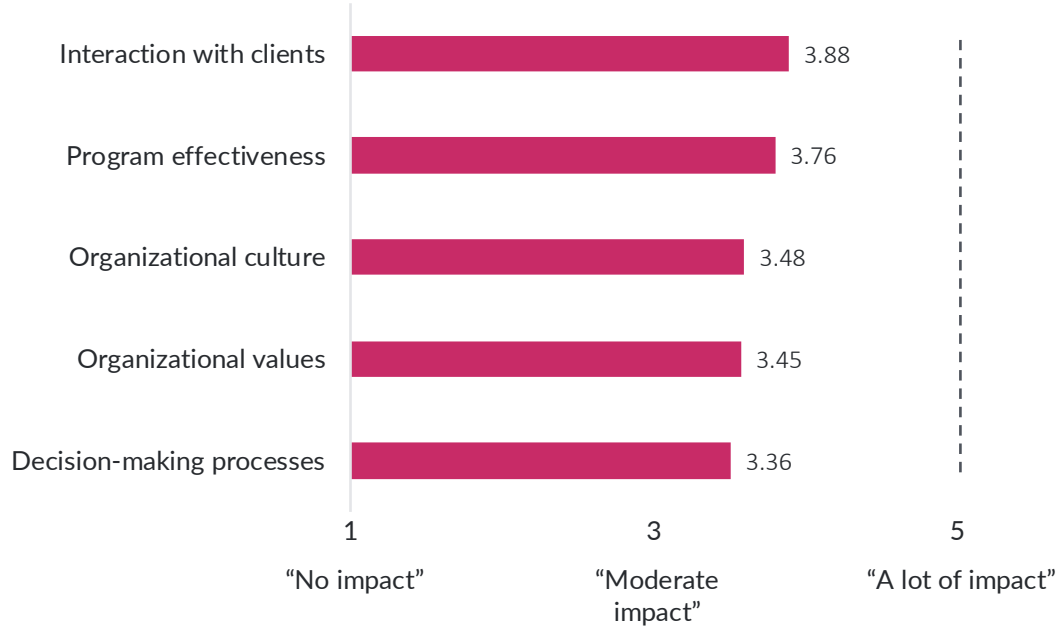
After 24 months with L4G, both program managers and agency leaders from the 2016 cohort have identified areas where L4G intersects with equity, diversity, and inclusion (EDI) and helps advance their work through feedback practice. The main way in which L4G has advanced EDI so far is by increasing client inclusion and ensuring that organizational decisions and programs respond to client needs. The data shows a few examples of how organizations are working on more equitable service provision by better understanding and addressing the needs of different types of clients, but the small number of examples suggests that organizations have not yet realized the full potential of this area of impact.

Feedback practice is also generating changes in organizational culture by elevating the value of client input, and organizations' intentionality in implementing inclusive practices. These practices can drive towards greater organizational effectiveness as they improve organizations' ability to understand, respond to, and meet client needs. Nevertheless, challenges related to EDI persist, particularly around translating surveys into clients' native languages using terms they can understand and administering surveys with hard-to-reach populations.

L4G improves organizations' ability to serve clients while changing organizational culture and values around listening to the people they seek to help.

Reflecting on their experience over two years of feedback practice through L4G, all 2016 cohort agency leaders at 24 months report that L4G has improved their organizations' ability to serve clients, with 58% indicating great improvement, and 42% indicating some improvement. As Figure 8 shows, organizations report that L4G has improved interaction with clients and program effectiveness, while also impacting organizational culture, values, and decision-making processes. Open-ended responses from program managers and agency leaders confirm this finding and provide examples of culture and practice changes, including how organizations share data internally with staff and an increased value for client feedback.

Figure 8 | Amount of impact L4G has on different areas within organizations[17]



[17] N = 33

All organizations plan to continue their feedback practice after their L4G grant.

One of the main indicators of the sustainability of feedback practice within organizations is their intention to continue collecting and using feedback after the L4G grant. At 24 months, all program managers and agency leaders report plans to continue collecting feedback after the grant, illustrating both their commitment to the process and the value of feedback loops.

There are some encouraging signs already. For example, after 12 months of engaging in feedback practice, 53% of organizations were surveying clients from only one program, while 48% were surveying clients from multiple programs in their organization. However, by the 24-month mark, 90% of agency leaders indicated that their organization had incorporated feedback collection in other areas of work. In open-ended responses, 24 agency leaders described that they had incorporated collecting and responding to feedback into other areas of work including internal use (9), and adding feedback to both existing (8) and new programs (3), among other activities (4). The data suggests that organizations with one-off interactions are more likely to use surveys across multiple programs than organizations that have indirect interactions. Similarly, organizations with one-off and sustained interactions are more likely to have incorporated feedback collection into other areas than indirect organizations.[18]

Nevertheless, the proportion of agency leaders who say they will continue but decrease the amount of feedback they collect has increased from 3% to 12% between the 12- and 24-month surveys, which may be a sign that organizations are starting to think about what the right amount of feedback is given their own needs and resources. There is some risk of organizations not continuing feedback practice and we will monitor the level of continued uptake.

L4G increases leadership and staff commitment to feedback practice.

Another possible indicator of sustainability is staff and leadership commitment to feedback practice. The data suggest that the 2017 cohort experienced significant increases in leadership and staff commitment to feedback practice at the six-month mark.[19] The data also show that program managers perceive leadership as significantly more committed than staff.[20] This finding is also true for the 2016 cohort, as the 24-month survey shows that program managers rate leadership commitment significantly higher than staff commitment.[21] Given that organizations from both cohorts have identified staff buy-in as an important element in making feedback practice successful in their open-ended comments, this finding could be a sign of a potential challenge for organizations.

---

[18] $p < .01$; the small number of "indirect" organizations in the sample may be a source of error in this test, and is thus an important limitation of the data.

[19] $p < .01$

[20] $p < .05$

[21] $p < .05$

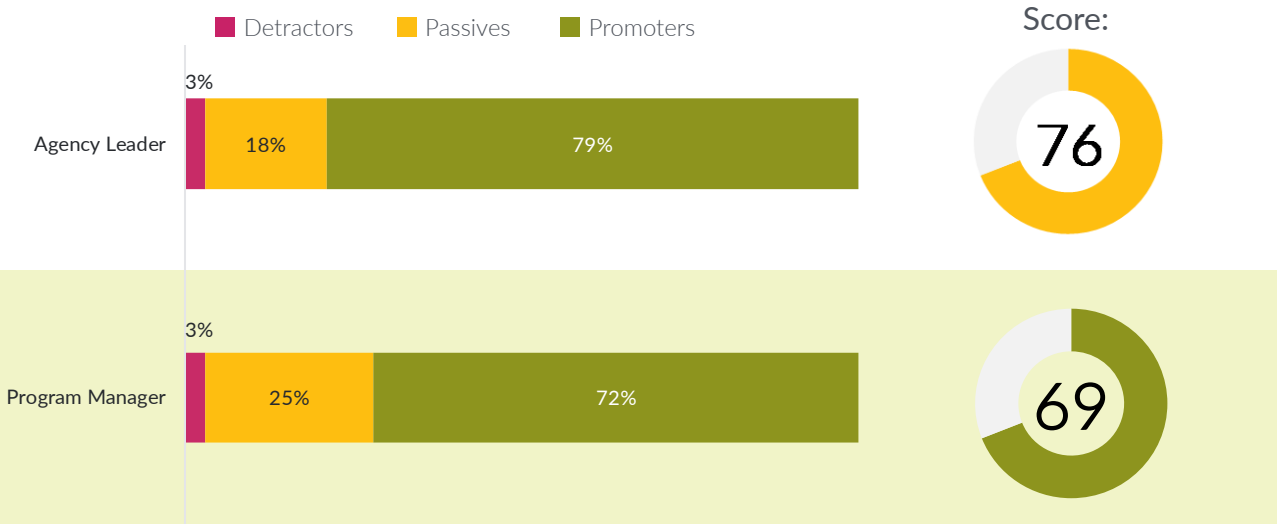
L4G's Support Model

This section examines how the L4G model has worked so far in supporting nonprofit partners' feedback practice. We consider how L4G has supported organizations overall and how specific elements and touchpoints within the model's design are working, illuminating both strengths and areas for potential improvement moving forward.

Organizations find the L4G model effective in supporting their feedback practice.

L4G has designed a support model to help nonprofits implement their feedback practice, which includes TA to guide organizations through each step of the process along with access to SurveyMonkey, a set of semi-standard survey questions, and sector-specific benchmarks. Organizations report high overall satisfaction with L4G using the same NPS® methodology that organizations use with their clients. The majority of program managers and agency leaders indicate that they would recommend L4G to other organizations (Figure 9).

Figure 9 | Agency leader and program manager Net Promoter Scores for L4G[22]



[22] Agency Leader N = 33; Program Manager N = 36. The NPS® methodology uses a 0-10 scale to measure how likely a respondent is to recommend an organization or service. Detractors are those who answer 0-6, passives are those between 7 and 8, and promoters are those between 9 and 10. The score subtracts the percentage of detractors from the percentage of promoters.

In the L4G model, the NPS® system includes two open-ended questions (compared to one question in the original system) that help illustrate L4G's strengths and opportunities for improvement. The majority of program managers and agency leaders recognize that L4G's main strength is providing technical support. Within the 34 comments about technical support, 20 mentioned general customer service and support, while 10 identified support in specific steps of the L4G process. On the other hand, both program managers and agency leaders identified improved training and sharing of learnings as the top two areas for potential improvement.

The amount of TA provided has remained adequate despite changes in the model, and organizations highly value the SurveyMonkey platform.

Overall, nonprofit partners remain highly satisfied with the amount of TA from the L4G staff despite the changes in the TA model across cohorts. Specifically, the majority of 2016 organizations rated both the amount of required engagement with L4G and TA support provided as being "just right" at the six-month mark. In the 12-month survey, we explored satisfaction with TA in greater detail by asking organizations to rate the different types of TA support provided, and the majority rated all types of support as "just right." Given the changes in the TA model between the 2016 and 2017 cohorts, we asked 2017 organizations to rate both the amount of one-on-one support and the amount of group support provided, and found high satisfaction with the amount of both methods of TA provision (Figure 10).

Figure 10 | Ratings of TA engagement and support across cohorts[23]

2016 Cohort

	6 months
Required engagement	92%
Support provided	94%

2017 Cohort

	6 months
Required engagement	87%
1 on 1 support provided	87%
Group support provided	88%

We also analyzed to what degree organizations in the 2017 cohort found TA support helpful, in light of which TA provider they were interacting with. The data show three statistically significant differences where organizations with one TA provider rated an item significantly more favorably than those working with another TA provider. Given that there were only three instances of statistically significant difference and that general ratings were favorable, this may not necessarily be an indicator of quality of service. However, the fact that there was some significant difference conveys the challenge of ensuring uniform, high-quality TA when more TA providers are involved, which is an important consideration for future L4G cohorts in the test rounds and public model.[24]

[23] 2016 6-month N = 36; 2017 6-month N = 58

[24] Significant findings were shared with the L4G team.

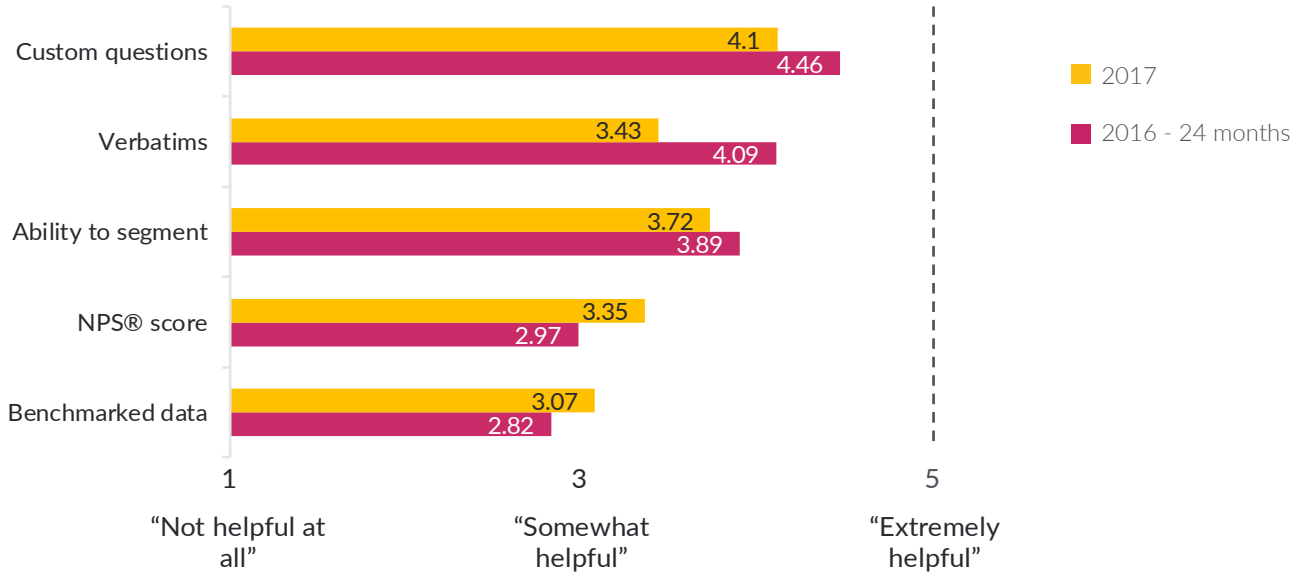
TA providers have different touchpoints with organizations at critical steps within the L4G process, so we asked organizations to rate each touchpoint in which they participated. All touchpoints received average helpfulness ratings between 4 and 5 (on a 1 through 5 scale) at six and 24 months, except for the "interpreting results webinar," which received a 3.37 score at 6 months. However, at 12 months, three out of six touchpoints received scores lower than 4, including the closing the loop call, the qualitative call, and the peer learning webinar. In addition, L4G has offered optional calls and webinars to further support organizations. Participation in these optional touchpoints ranged from 43% to 60% among the 2016 cohort, and program managers rated them as highly helpful (4.05 and above on a 1 through 5 scale).

Another added value for organizations that participate in the L4G model is having access to SurveyMonkey to facilitate data collection and reviewing results. Overall, both cohorts indicate high levels of satisfaction with SurveyMonkey's digital platform in their six-month surveys. However, 2017 organizations rated their experience reviewing results through SurveyMonkey significantly higher than 2016 organizations.[25] L4G staff hypothesize that the 2017 cohort had more previous exposure to the SurveyMonkey platform; however, this survey data does not allow us to confirm that as the reason for the higher rating.

Organizations experience some elements of the L4G surveys as more helpful than others.

In terms of the survey design, organizations find that custom questions are most helpful for generating actionable data, followed by verbatims and ability to segment data (Figure 11). Verbatims were significantly more helpful for 2016 than for 2017 organizations.[26] A possible explanation for this difference is that data may appear more valuable or actionable in retrospect because organizations have had time to make changes based on the data.

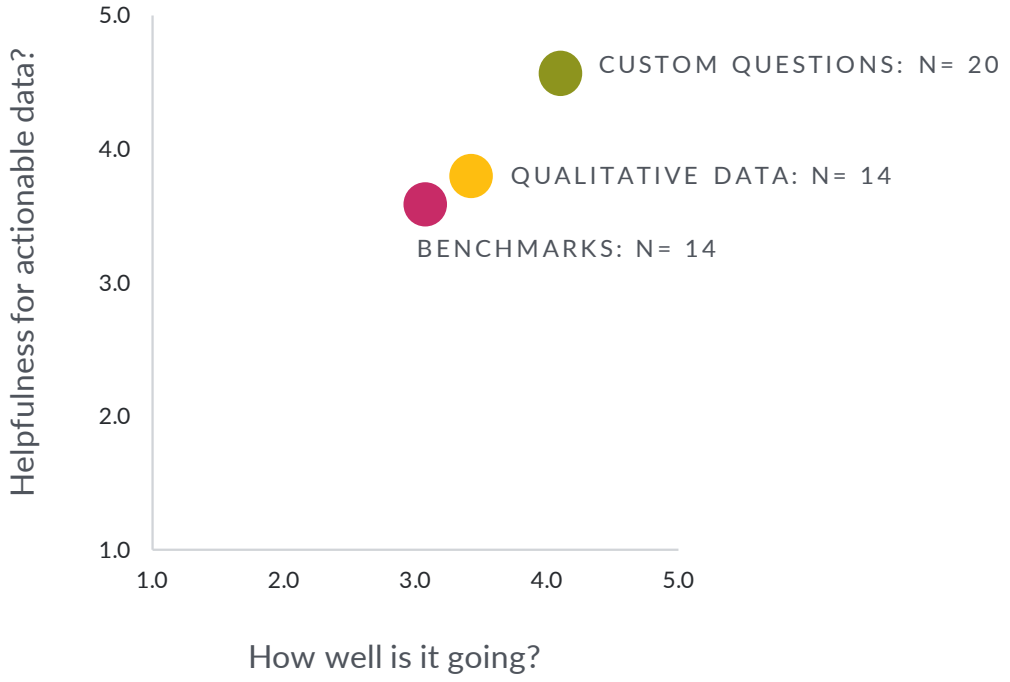
Figure 11 | Helpfulness of survey elements in generating actionable data[27]



[25] $p < .01$
[26] $p < .05$
[27] 2016 24-month N = 38; 2017 6-month N = 58

Given that the NPS® and benchmarks are core elements of L4G's value proposition for organizations, and that they are the least helpful elements for generating actionable data, we explored the lower ratings in greater detail. Specifically, we found that benchmarks are the lowest rated items in how helpful they are and how well organizations' experience with them is going so far (Figure 12). This is an emergent finding that we will continue to monitor in upcoming analyses.

Figure 12 | Helpfulness and of ease of use of survey elements for the 2017 Cohort[28]



Open-ended data provides some insight into what organizations find challenging about benchmarks. Questions and concerns about how benchmarks are designed, how they apply to different types of organizations, and whether they have been developed enough for the nonprofit sector surface in both the 2016 and 2017 cohort surveys. In addition, some 2017 organizations expressed a lack of knowledge about what benchmarks are and how to use them.

There may be a correlation between organizations reporting that elements are working well and the extent to which they find them helpful in generating actionable data. The implication is that organizations' experience with survey elements depends on each element's realized potential to yield actionable data, and not necessarily on its hypothetical potential. Therefore, TA support in fulfilling the potential of NPS® and benchmarks may be especially important to increase the overall value organizations get from these core elements of L4G.

[28] 2017 six-month N = 58. The way we collected data did not allow us to analyze the NPS® or other elements in this way, but we will adjust in future surveys to conduct the full analysis. Future surveys will also provide a larger sample size for the analysis.

There are common successes and challenges throughout the steps that both L4G and new organizations can learn from as they move feedback practice forward.

When considering specific steps in the process, the 2017 cohort mentioned that within the survey design step, developing custom questions, using the survey templates with standardized questions, and the interactions with TA providers were aspects that were going particularly well. Similarly, within the survey administration step, flexibility in choosing an appropriate method to administer the survey, when and where they administered the survey, and client understanding of survey questions were going well. Finally, using SurveyMonkey and quantitative data analysis and interpretation were successful aspects of the interpreting results step. On the other hand, organizations from both cohorts have identified similar challenges they face throughout the process, summarized in figure 13.

Figure 13 | Challenges at different steps in the L4G feedback process

Survey Design	Administration	Interpretation	Responding to Feedback	Closing the Loop
• Developing custom questions. • Creating questions that are clear for clients. • Integrating NPS® into their survey design.	• Low response rates. • Unique constituent needs (language, timing, etc.). • Clients having difficulties using tablets. • Capacity for survey administration. • Administering surveys in some settings.	• Finding actionable data. • Small sample size. • Not enough resources for interpreting results. • Not enough in-depth qualitative comments. • Not being able to respond given the "one-off" nature of feedback received.	• Lack of capacity (time and staffing), and resources (infrastructure, funding) to respond to client feedback.	• Finding the appropriate way to close the loop. • Finding the right time to close the loop. • Bringing along staff and crafting a message before closing the loop.

Looking Ahead

The findings in this report have implications for Shared Insight and L4G as they move forward in their strategy toward a public, web-based L4G model, and for ongoing evaluation and learning. Implications for the test rounds and the public model center around what L4G can learn from this iterative process to better adapt the model before it opens to the public, and how L4G can best position itself to add value to organizations' work through a less controlled and supported system.

Focus support in generating insights and increasing ability to increase program effectiveness.

As this report highlights, the extent to which organizations gain insights or increase technical abilities affects the extent to which L4G impacts program effectiveness. Therefore, L4G can design supports and touchpoints around these two factors, considering how to best leverage web-based products to continue to deliver high-quality support that makes the biggest difference for organizations. Knowing that TA has been particularly helpful in areas like quantitative analysis support, L4G can take these more successful elements forward while also adapting and improving the more challenging areas. For instance, because organizations seem to have difficulties using the NPS® system, benchmarks, and qualitative data in their analysis, L4G could better position TA as a way to better use that data to generate actionable insights. It is important to note that these difficulties are an emergent finding, and that L4G staff has already taken steps through TA to address these issues. We will continue to monitor to what extent those actions are responding to these potential trip wires in nonprofits' feedback practice.

The right thing, the smart thing, the transformative thing – for whom?

The data from these two cohorts suggests that smaller organizations and those with sustained client interactions are gaining more insights from L4G. These findings suggest that L4G could market both test round participation and the public model to organizations with those characteristics, as they are more likely to benefit from the experience and thus may be more willing to pay for the service. However, this is still an emergent finding that we will continue to explore. Further research could also help understand whether there are certain upfront requirements or conditions that organizations should meet before engaging in L4G.

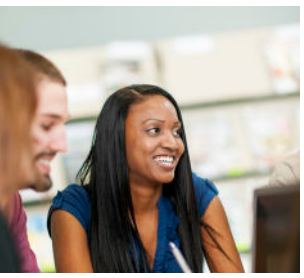
L4G Evaluation Moving Forward

As L4G and Shared Insight drive feedback practice forward through the L4G initiative, ORS Impact will continue to be a thought partner committed to supporting data-informed learning and decision making, holding up the mirror to the work, and remaining adaptable in its evaluation practices. Over the next three years, we will focus our attention on five specific areas related to L4G and nonprofit feedback practice, including questions like:

1. **Nonprofit feedback outcomes given the changes in L4G's model:** How do model changes affect the level of uptake and organizational outcome achievement, and why?
2. **Value of L4G/feedback practice:** What do nonprofits find valuable about participating in L4G, and why and what types of nonprofits find more value in L4G? How do nonprofits use feedback, monitoring, and/or evaluation as part of the "three-legged stool" model in their organizations?
3. **Sustainability of feedback practice:** To what degree do former nonprofit partners continue with feedback practices and what methods do they use? To what extent do former nonprofit partners follow through on planned changes or maintain changes made based on L4G insights?
4. **Client experience:** How do L4G nonprofit partner clients experience the surveys? How has feedback changed their experience with the organization? Do client perceptions differ by race, age, gender, or other characteristics?
5. **Feedback as an equity tool:** To what degree does feedback build or augment L4G nonprofit organizations' own work to address issues of equity, diversity, or inclusion? If, when, and how does feedback change organizational practices that lead to greater equity, diversity, or inclusion internally?

Closing Thoughts

L4G was created to be a mechanism to support nonprofits in listening to the people at the heart of their work while building infrastructure, engaging funders, and capturing and sharing lessons learned to positively catalyze the feedback practice field. During the first two cohorts, L4G has successfully supported 112 nonprofit partners in their feedback practice, increasing their ability to listen to clients, gain insights, and make changes in policies, practices, and programs to better serve their clients. L4G has also impacted organizational culture and advanced equity, diversity, and inclusion by elevating the value and importance of client feedback as a mechanism to shift power dynamics and provide clients the opportunity to shape the services that benefit their communities. The experience so far demonstrates many successful aspects of the L4G model and identifies areas of potential improvement as L4G prepares to expand its impact in the feedback practice field.



ORSIMPACT
1100 OLIVE WAY, SUITE 1350
SEATTLE, WA | 206-728-0474
ORSIMPACT.COM

Appendix A: Methodology, Data, and Results

Methodology

ORS Impact has conducted surveys with nonprofit partners in the 2016 and 2017 Listen for Good cohorts for our evaluation of Listen for Good. We surveyed program managers and agency leaders from the 2016 cohort at three time points: six, 12, and 24 months into their grant cycles. Thus far, we have collected 6-month survey data from program managers of the 2017 cohort.[1] Together, these surveys resulted in seven datasets that we merged to create one quantitative dataset.

Having one dataset with all seven surveys allowed for several different statistical analyses. First, we conducted statistical comparisons *within organizations*, including comparisons between agency leaders and program managers as well as comparisons between different time points. We also made statistical comparisons *between organizations*, conducting segmentation analyses by budget size, the organization's area of focus, and the nature of the organization's relationship with their beneficiaries. Finally, we made statistical comparisons *between cohorts*, examining how data from the 2017 cohort compares to data from the 2016 cohort.

We prepared the data for statistical analyses by recoding grouping variables (e.g., budget size) into categories to run comparative analyses and creating scale scores. For instance, we used multiple survey items to create scale scores for: (a) organizational abilities, (b) the degree to which new insights were drawn, and (c) perceived internal impacts of the Listen for Good Initiative. We used these scale scores in our descriptive and inferential statistical analyses. As listed in Figure 14 below, these analyses included chi-square tests, paired samples *t* tests, independent samples *t* tests, one-way ANOVAs, hierarchical linear regressions, Mauchley's test of sphericity, and binary logistic regression. For each of these analyses, we reported findings as statistically significant when $p < .05$ and trending toward significance when $p < .1$. Additional output from all tests is available upon request.

For qualitative data from the surveys, we used a thematic analysis approach to code the data for prevalent themes across responses for each individual question, calculated the frequency with which comments reflected each theme. Finally, we looked searched for similarities and differences in how organizations spoke about prevalent themes both within and across cohorts.

[1] With the second cohort of Listen for Good, we decided to only collect data from the Program Manager and will only collect data at two time points (six months and 12 months) to decrease overall burden given the deep lessons learned from the first cohort.

Inferential Analyses and Data Sources

Figure 14 | Statistical Analyses and Data Sources

Comparison	Analysis Used	Data Source Used
Ability Over Time	Paired samples *t* Test	All data sources used
Equality of Variability in Ability Scores Between Time Points	Mauchley's Test	2016 Cohort, Agency Leader and Program Manager Data at 6, 12, and 24 Months
Ability and the Likelihood of Seeing Impact on Program Effectiveness	Hierarchical Linear Regression	2016 Cohort Program Manager Data at 24 Months
Organization's Interactions with Clients and Ability Scores	Hierarchical Linear Regression	2016 Cohort Program Manager Data at 24 Months
Insights Over Time	Paired Samples *t* Test	2016 Cohort, Program Manager Data at 12 and 24 Months
Insights and Likelihood of Seeing Impact on Program Effectiveness	Hierarchical Linear Regression	2016 Cohort Program Manager Data at 24 Months
Insights by Organization's Budget Size and Interactions with Clients	Hierarchical Linear Regression; Chi-Sq. Test	2016 Cohort Program Manager Data at 24 Months
Commitment over time; Commitment within organizations	Paired Samples *t* Test	2016 and 2017 Cohorts, Program Manager Data at 6, 12, and 24 Months
Commitment by Cohort	Independent Samples *t* Test	2016 and 2017 Cohorts, Program Manager Data at 6 Months
Ratings of Helpfulness by TA Provider	One-Way ANOVA	2017 Cohort, Program Manager Data at 6 Months
Helpfulness of Survey Monkey by Cohort	Independent Samples *t* Test	2016 and 2017 Cohorts, Program Manager 6-Month Data

Hierarchical Linear Regression Results

Figures 15 and 16 show the output from hierarchical linear regression analyses, which yielded the statistically significant results mentioned in the report's findings. Additional output for other tests is available upon request.

Figure 15 includes the results from a regression analysis with the outcome as a 1-item measure of the extent to which 2016 cohort program managers at 24 months believed that their involvement in Listen for Good had an impact on program effectiveness. Model 1 served as a baseline regression model to predict the outcome based solely on organizational characteristics. These included budget (compared to the smallest budget group, < $1M), the nature of the organization's interaction with clients (compared to organizations with sustained interactions), and the organization's area of focus (compared to those focused on Economic and Community Development). In Model 2, we added program manager ratings of the organization's ability related to feedback practices as an additional predictor in the model. Organizational ability was measured as the average of seven survey questions related to feedback abilities (e.g., ability to implement surveys, achieve a high response rate, analyze and interpret data). These seven items demonstrated sufficient internal consistency (α = .81). Lastly, in Model 3, we added program manager insights as the final predictor in the model. Program manager insights gained was measured as the average of five items (e.g., new insights in client needs and preferences, interactions), again demonstrating sufficient internal consistency (α = .72).

Figure 15 | Regression Output for Analysis of Predictors of Likelihood of Seeing an Impact on Program Effectiveness

Model 1			Model 2			Model 3		
Predictor	*B*	*SD*	Predictor	*B*	*SD*	Predictor	*B*	*SD*
(Constant)	3.82***	0.36	(Constant)	1.41	1.18	(Constant)	-0.63	1.51
Budget (Compared to < $1M)			Budget (Compared to < $1M)			Budget (Compared to < $1M)		
1M to 5M	-0.24	0.37	1M to 5M	-0.25	0.34	1M to 5M	0.13	0.38
5M to 10M	-0.57	0.47	5M to 10M	-0.75	0.45	5M to 10M	-0.36	0.47
More than 10M	-0.58	0.37	More than 10M	-0.78**	0.36	More than 10M	-0.59	0.35
Nature of Interaction (Compared to Sustained)			Nature of Interaction (Compared to Sustained)			Nature of Interaction (Compared to Sustained)		
Indirect	0.55	0.46	Indirect	0.33	0.44	Indirect	0.40	0.42
One-Off	-0.58*	0.29	One-Off	-0.30	0.30	One-Off	0.08	0.34
Issue Area (Compared to Economic and Community Development)			Issue Area (Compared to Economic and Community Development)			Issue Area (Compared to Economic and Community Development)		
Education	-0.96**	0.46	Education	-0.60	0.46	Education	-0.40	0.45
Health	-0.08	0.39	Health	0.05	0.37	Health	0.30	0.37
Other	0.59	0.37	Other	0.54	0.35	Other	0.49	0.33
Human Services	0.84**	0.38	Human Services	0.67**	0.36	Human Services	0.51	0.35
			Ability	0.59**	0.28	Ability	0.64**	0.26
						Insights	0.55*	0.28

* $p < .10$, ** $p < .05$, *** $p < .01$

Figure 16, shows results from a similar process, this time with the outcome being a 5-item measure of the extent to which 2016 cohort program managers at 24 months felt that they gained insights as a result of participating in Listen for Good. Model 1 again served as a baseline regression model using only the organizational characteristics described above in reference to Figure 15. Model 2 in this analysis included the same predictors used in Model 1, plus program manager ratings of the organization's ability related to feedback practices (see above for description).

Figure 16 | Regression Output for Analysis of Predictors of Gaining Insights

Model 1			Model 2		
Predictor	*B*	*SD*	Predictor	*B*	*SD*
(Constant)	3.35***	0.23	(Constant)	3.70***	0.82
Budget (Compared to < $1M)			Budget (Compared to < $1M)		
1M to 5M	-0.69**	0.24	1M to 5M	-0.69**	0.24
5M to 10M	-0.73**	0.30	5M to 10M	-0.71*	0.31
More than 10M	-0.38	0.24	More than 10M	-0.35	0.25
Nature of Interaction (Compared to Sustained)			Nature of Interaction (Compared to Sustained)		
Indirect	-0.18	0.29	Indirect	-0.14	0.31
One-Off	-0.65**	0.19	One-Off	-0.69**	0.21
Issue Area (Compared to Economic and Community Development)			Issue Area (Compared to Economic and Community Development)		
Education	-0.32	0.29	Education	-0.37	0.32
Health	-0.44	0.25	Health	-0.46*	0.26
Other	0.09*	0.24	Other	0.10	0.24
Human Services	0.28	0.24	Human Services	0.30	0.25
			Ability	-0.09	0.19

* $p < .10$, ** $p < .05$, *** $p < .01$