Connection and Reflection: Peer-led Conversations as an Evaluation Approach with Young People

Jennifer Jocz, *Educational Development Center* iD
Laycca Umer, *New York Hall of Science* iD
Katherine McMIllan Culp, *New York Hall of Science* iD

Abstract: This paper examines the use of peer-led conversations as a culturally responsive evaluation method in a youth-focused out-of-school time (OST) program. Drawing on principles of culturally responsive evaluation and youth participatory action research, the study explores whether and how peer-led reflection sessions can address common challenges in OST evaluation, including power imbalances, limited evaluator–youth relationships, and restricted time for data collection. The program evaluation, through collaboration with program staff and youth fellows at the New York Hall of Science, incorporated three peer-facilitated reflection conversations. Findings suggest that peer-led conversations can provide a comfortable space for youth to share experiences, offer opportunities for connection and self-reflection, and yield valuable feedback to inform program improvement. The study concludes with recommendations for evaluators and practitioners seeking to incorporate more inclusive, youth-centered evaluation strategies in OST settings.

Keywords: Peer-led conversations, Youth participatory evaluation, culturally-responsive evaluation

Introduction

One role of an evaluator is to gather information about the thoughts and experiences of individuals impacted by an intervention and to use the findings to examine the effectiveness of the intervention and inform improvement. If evaluations are meant to determine the extent to which something has value, then it is imperative to situate that in the values and beliefs of the culture—both the identities and lived experiences of participants and the norms and practices of the program setting—in which the program is occurring and for whom it is meant to serve.

Culturally Responsive Evaluation

Culturally Responsive Evaluation (CRE) is an inclusive framework that recognizes that evaluations are not devoid of values and beliefs. Rather, it encourages evaluators to reflect on their own positionality and

values while also valuing the culture and context of the programs, communities, and individuals they are working with (Acree & Chouinard, 2020; Hood et al., 2015; Hook et al., 2013).

In this study, we define "culture" broadly to encompass both the identities, experiences, and community norms youth bring with them, and the interactional norms, values, and practices embedded in the program setting. A culturally responsive approach attends to both, engaging youth in ways they identify as meaningful and relevant, while aligning with the specific context and goals of the program (SenGupta et al., 2004).

Because of this, CRE does not follow a one-size-fits-all approach, it is not something that can be prescribed or standardized. CRE requires the capacity to understand and appreciate other cultures, to be cognizant of one's own biases, and the flexibility to respond and adapt the evaluation based on one's awareness of oneself and others.

In their review of the literature, Acree & Chouinard (2020) found that culturally responsive evaluations generally intend to be empowering and inclusive, engaging voices that are traditionally left unheard or underheard in evaluations such as those of youth or individuals from non-dominant cultural groups. Furthermore, culturally responsive evaluations tend to be more collaborative and participatory, which can increase individual stakeholders' evaluative skills and positive attitudes towards evaluation. However, even in culturally responsive evaluations, evaluators make many decisions about, and therefore retain power related to, how evaluations are conducted and how results are interpreted (Acree & Chouinard, 2020). Additionally, evaluators need to be careful to integrate culturally responsive methods, not appropriate them (Brewington & Hall, 2018). Evaluators should develop their evaluation in collaboration with stakeholders from the cultures and communities in which the evaluation is situated to determine how culturally responsive methods can be utilized in practice.

Power and Positionality in Evaluation

Evaluation is not neutral; it is inherently shaped by power dynamics. Evaluators determine what questions are asked, whose voices are included, how data are interpreted, and what recommendations are made. This control places evaluators in a position of authority, often leading to unintentional reinforcement of existing hierarchies and inequities. The dominance of Western scientific traditions and standard methodologies can marginalize other ways of knowing and devalue local and cultural knowledge (Mertens, 1999; Thomas & Campbell, 2021).

These power differentials are especially pronounced in evaluations that engage with historically marginalized communities. When evaluators are not reflective about their own positions or fail to build equitable partnerships, evaluation processes can become extractive, benefiting institutions while offering little to those whose experiences are being studied (Rodriguez & Acree, 2021).

In youth-serving programs, these dynamics are even more amplified. Youth have historically been viewed as a vulnerable population given little, if any, chance to speak out about programs and practices that impact them (Schelbe et al., 2015). This exclusion silences their voices and ignores their social and cultural capital. It removes their power to express their realities and inform programs and policies that impact them. This can result in interventions being developed without the input of the youth they are meant to serve.

Traditionally, research about youth experiences, including program evaluation, relied on adult proxies (e.g., caregivers, teachers) because adults were assumed to be more reliable informants (Schelbe et al., 2015). This practice reflects broader sociohistorical constructions of "youth" and "adolescence" as distinct—and often transitional—categories that have been shaped within Euro-American paradigms, including post–World War II consumer culture and Cold War-era concerns about youth, risk, and maturity (Lesko, 2012; Savage, 2007). Because these frameworks have frequently been generalized as universal, they can obscure

the cultural and historical variability of what counts as "youth," and they can position young people as "not yet" credible knowers of their own lives.

Against this backdrop, constructivist and participatory approaches emphasize that young people actively interpret and shape their social worlds and can therefore provide credible, situated insights about their own experiences (Kirk, 2007; Mason & Hood, 2011). Framed this way, the central question becomes: Can youth provide credible insights about their own experiences—and what conditions make those insights most visible and actionable in evaluation? Rights-based frameworks similarly affirm young people's entitlement to express their views in matters that affect them (United Nations, 1989). In practice, evaluation approaches that reduce adult–youth power asymmetries and build from shared experience—such as peer-led reflection—offer one feasible way to operationalize this stance in OST contexts.

Participatory Action Research and Evaluation (PARE)

In the late 1990s, the evaluation field began pushing for the inclusion of program participants in the evaluation, especially participants from marginalized groups. Early efforts toward inclusion often involved participants serving as consultants or reviewers at specific points in the process, while evaluators retained primary control over design and interpretation. This differs from contemporary participatory approaches, which more often emphasize co-design, shared decision-making, and participant leadership throughout the evaluation process (Schelbe et al., 2015). Evaluators took the viewpoint that, without these voices, the evaluation would be incomplete and lack credibility. These early inclusion efforts marked a shift from traditional, evaluator-led approaches, but they were generally more consultative than collaborative by today's standards.

Inclusive evaluations strengthen the rigor and validity of the work (Mertens, 1999). A review of the literature suggests that participatory methods can help move research forward by balancing scientific standards (by incorporating the expertise of the researcher or evaluator) with cultural and social validity (by including the voices of participants) (Cargo & Mercer, 2008).

In recent years, Youth Participatory Research and Evaluation (YPARE) has been used as a way to allow young people to have a voice in evaluating the programs that serve them (Sabo-Flores, 2007). The extent of youth engagement can range from passive "consultants" who give feedback but do not actively participate in the implementation of the evaluation, to fully engaged "leaders" of evaluations who are responsible for designing and carrying out their own evaluation activities (Richards-Schuster & Elliott, 2019).

In general, involving youth in research and evaluation practices has been shown to have positive influences on youth. This includes growing their social and cognitive skills, increasing agency, supporting academic and career pathways, helping to build intergenerational relationships, providing leadership opportunities, and increasing civic engagement (Anyon et al, 2018; Powers & Tiffany, 2006).

However, implementing YPARE is not without its challenges. It requires the commitment of adults and organizations to authentically include youth and treat them as valued team members (e.g., providing monetary compensation, supporting them in presenting their findings to others) to support stained involvement (Fox & Cater, 2011). Furthermore, successful YPARE requires a large time and resource commitment and the sustained engagement of youth over the course of the study. Suleiman et al. (2006), reflecting on their experiences with youth-led research projects, note that youth had limited time to attend meetings and participate in research activities leading to missed deadlines and failure to meet their goals.

Issues of validity and reliability also challenge youth participation in research and evaluation. Historically, validity and reliability focused on traditional scientific criteria such as detailed documentation of the research process, triangulation, and internal checks (e.g., inter-rater reliability) (Lewis, 2009). These measurements of quality and trustworthiness may not fully capture or reflect the goals of participatory methods, which focus

on empowerment and inclusion (Kim, 2016). When viewed through the traditional lens of validity and reliability—focused on criteria such as detailed documentation, triangulation, and inter-rater reliability, participatory approaches may appear to fall short (Lewis, 2009). This perspective overlooks other benefits of these methods, including forms of validity such as cultural and social validity, which participatory approaches are well-positioned to promote (Cargo & Mercer, 2008).

At the same time, reliability in youth-centered participatory work is often best understood as situated dependability rather than as fixed reproducibility across time and setting. Youth perspectives may shift as relationships evolve, program phases change, or external conditions shape what feels salient. Accordingly, consensus achieved through participatory processes may be fluid. Evaluators can strengthen reliability by (a) documenting contextual conditions (who participated, when, where, and what was happening in the program/community), (b) using iterative cycles or repeated conversations to examine which themes persist, transform, or diverge over time, and (c) triangulating youth-generated insights with observations, artifacts, or other stakeholder perspectives. Importantly, when youth views change over time, that movement can be interpreted as a meaningful finding about developmental and contextual dynamics.

Despite these challenges, YPARE offers a compelling path forward for those seeking to engage youth in ways that are culturally responsive, equitable, and transformative Mason & Hood, 2011). However, researchers and evaluators wishing to implement YPARE should take into consideration youths' needs and interests, especially when time or resources may be insufficient to build youths' skills and support their integration into the project.

While the peer-led reflection conversations in this study were not full YPARE, they moved beyond early consultative models by giving youth a central role in shaping and leading evaluation activities within their shared program context. This approach was particularly well-suited to the realities of the program in which it was implemented, where time, staffing, and evaluation resources were limited, and where creating a comfortable environment for youth voice was essential.

Evaluating Out-of-School Time Programs

Evaluating youth programs in out-of-school time (OST) settings comes with a unique set of challenges. Fully participatory approaches, while valuable, can be difficult to sustain in these contexts due to constraints in time, staffing, and relationship-building. First, youth attendance in OST programs can be sporadic, making it difficult to track growth or outcomes over time. Furthermore, OST programs often prioritize fun, engagement, and enrichment, making it difficult to find time for evaluation activities that do not pull participants away from their program. Additionally, traditional data collection methods such as surveys and adult-led interviews can feel especially out of place in a highly interactive OST setting.

Finally, determining the appropriate person to carry out the evaluation can be challenging. Internally, OST staff are unlikely to be trained in data collection or evaluation. They may also prefer to stay solely in their roles as program facilitators to avoid creating additional power imbalances between themselves and program participants. When utilizing external evaluators, there is generally not enough time or funding to allow evaluators to connect with youth before data collection takes place. Without this foundation, youth may feel reluctant to share their experiences with the evaluator.

This paper focuses on the development and use of peer-led reflection conversations in an OST setting as a way to involve youth in the evaluation activities while mitigating the challenges of time and power imbalances. In this sense, peer-led conversations offered a practical middle ground — retaining many of the benefits of participatory evaluation, such as youth voice, relevance, and shared ownership (Cargo & Mercer, 2008), while remaining feasible within the realities of an OST environment.

Methodology

Building on the need for an approach that balanced the benefits of participatory evaluation with the practical constraints of OST settings, this study implemented peer-led reflection conversations as a feasible and context-appropriate middle ground.

Participants

Participants included program staff and young people from the Science Career Ladder (SCL) program at the New York Hall of Science (NYSCI). According to their website (www.nysci.org), "located in Queens – America's most diverse county – NYSCI is committed to creating a world where diversity unlocks innovations, and where people learn to use science, technology, engineering and math to tackle complex 21st century challenges."

The SCL program aims to support high school and college students' interests and access to STEM careers and college pathways. The SCL's flagship program employs young people (called "Explainers") to work on the museum floor while also offering ongoing training, learning opportunities, and mentorship. Explainers begin the program as "trainees" and can move up the rungs of the ladder over time as they develop their skills and expertise to become "Program Explainers" and "Senior Explainers", positions with increasing levels of responsibility. Since 1986, more than 3,500 young people have participated in the SCL program. The SCL program engages many youth that are typically underrepresented in STEM with approximately 60% of participants identifying as female and 85% coming from minority backgrounds. Additionally, the demographics of SCL participants reflects the diversity of the local community with 80% of Explainers residing in Queens.

This study examines youth experiences in a public health-related fellowship which is part of a larger opportunity for Explainers in the SCL program. This fellowship engaged a cohort of six youth fellows to co-design exhibit experiences with NYSCI staff. They met twice a week, sometimes virtually and sometimes in person, from December 2023 to June 2024. The goals of the fellowship included increasing the Explainer's awareness and understanding of public health and public health-related careers and providing them with knowledge and skills that they can apply to their future education and career pathways.

Engaging OST Staff and Participants in Evaluation Design

Conversations with NYSCI staff and fellows informed the design of the study.

During individual baseline interviews all fellows described a preference for conversations (either one-on-one or in a group) as opposed to providing written feedback. They noted that talking, especially as a peer group, helps them express their thoughts more easily. They also offered suggestions for designing the conversations including grounding questions in a common experience, organizing conversations so that participants are in groups with others they are familiar with.

Senior Explainers proposed utilizing the Explainer conversation model. This is a model used at NYSCI that trains Explainers to facilitate conversations with their peers around topics that are relevant to them (e.g., mental health and bullying).

NYSCI staff on the fellowship team were consulted to discuss how feedback from the Explainers could be utilized in the evaluation. All team members appreciated the idea of utilizing peer-led conversations as a means of collecting feedback. These peer-led conversations could be incorporated into the fellowship program as a group reflection activity so the Fellows view it as part of the experience. Fellows were presented with this idea and agreed that they would like to take part in the peer-led conversations.

Questions for each conversation were developed in collaboration with NYSCI staff members and with input, when possible, from the fellows. Evaluators also sought feedback from fellows after each conversation

to inform the design of the next. These collaborative dynamics between program staff and the evaluator were part of the broader system of relationships and power dynamics that shaped what forms of youth

Implementation of Peer-led Conversations

To inform the program evaluation, fellows participated in three reflective conversations, each led by one of their peers. Fellows took turns facilitating the conversations, with each conversation consisting of four questions. Questions centered on fellows' experiences in the program, including successes, challenges, and questions.

Each session lasted about 30 minutes and took place in person once a month for three months, serving as dedicated reflection time for the fellows. They met in a private room without program facilitators present, and the sessions were not recorded to protect privacy and encourage open conversation. The evaluator joined via Zoom, explaining at the outset that their role was to take notes and provide assistance if questions arose.

After each conversation, the evaluator provided formative feedback to the NYSCI team to highlight areas where fellows had questions or needed additional support. To inform program improvement and capture fellows' perspectives on the peer-led model, a summative report was developed based on notes from the conversations as well as baseline and end-of-program interviews.

Results

Fellows' Views of Peer-led Conversations

Notes from peer-led conversations and individual interviews were coded using emergent coding and analyzed for themes. These conversations not only revealed patterns in fellows' experiences but, in some cases, also led to immediate program changes within the same fellowship cycle (described later in this section). After the first and second sessions, fellows expressed a desire to continue meeting in this way, indicating that they found the conversations both meaningful and beneficial. The specific benefits they identified are detailed in the sections that follow.

Comfort and Shared Experience

Fellows shared that the peer-led conversations felt comfortable because they knew one another and understood each other's experiences. They were working toward the same goals, facing similar challenges, and could speak candidly without the presence of program facilitators. Talking to someone outside of the program, one fellow noted, would feel more like "giving a presentation."

The conversations also provided a space away from NYSCI staff to connect and talk openly about their experiences. As one fellow put it, "No matter how much [NYSCI staff] try, they can never be as comfortable to talk to about specific issues." Another explained, "When we have sessions with Fellows, we can open up about ideas and concerns with the program. It serves as a reflection time, but you're more comfortable talking about experiences."

Although we did not directly gather data on fellows' expectations going into the conversations, program staff reported making changes in response to feedback shared during these sessions and discussing those changes with the group. These visible, in-the-moment adjustments likely signaled to fellows that their input was valued and acted upon, potentially strengthening their sense that the conversations could lead to meaningful improvements in the program.

Immediate Program Changes

Two notable examples illustrate how fellows' feedback during the peer-led conversations translated into program adjustments during the same fellowship cycle. Midway through the program, fellows expressed wanting a clearer understanding of how their exhibits connected to NYSCI's broader institutional goals. In response, staff spent additional time reviewing these goals with the group. Fellows also shared a need for more direction and support during the development process. This led the program team to recruit additional fellows to help share the workload and to provide more regular and explicit feedback throughout exhibit development. Implementing and openly discussing these changes may have reinforced for fellows that their contributions during the peer-led conversations had a tangible impact on the program.

Opportunities for Connection and Reflection

Fellows noted that it was valuable to learn what others were working on because they "didn't hear a lot from the other Fellows—they're all busy building their own exhibits.". These conversations created space to reconnect, share updates, and learn from one another's challenges and strategies.

Additionally, fellows viewed these sessions as an opportunity to reflect both individually and collectively. The peer setting provided a low-pressure environment which encouraged ongoing reflection about fellows' experiences, areas of confusion, and progress towards goals. One fellow described, "It's not just to talk to someone about something, but also to understand your own self."

Turns of Talk Per Question

In order to determine the extent to which there was back-and-forth discussion in response to the discussion questions, the minimum, maximum, and average number of turns of talk per question were calculated. Turns of talk are a common unit in interaction analysis and can serve as a coarse indicator of interactional density and participation balance across speakers (Sacks, Schegloff, & Jefferson, 1974; Knight, Wise, & Sieke, 2016). We interpret turn counts as descriptive rather than as a direct measure of conversational quality. Accordingly, we consider these counts alongside the coding of response types and emergent themes described in subsequent sections.

As shown in Table 1, overall, there was an average of 7.8 turns of talk per question across all sessions. The minimum was three turns of talk per question with a maximum of 24 turns of talk. Three fellows attended each peer-led conversation. Therefore, conversation three had one question in which there was no back-and-forth conversation; however, all other questions led to fellows responding to one another and deepening the conversation.

Table 1. Turns of talk per question

Conversation Number	Average # turns of talk/ question	Min-Max turns of talk/ question
1	12.8	5-24
2	6.2	5-9
3	4.5	3-6
Overall	7.8	3-24

The Nature of Responses

The nature of individual responses during peer-led conversations was coded in two ways: (1) the relationship of the response to the overall conversation, and (2) emergent themes. Specific codes are shown in Table 2.

Table 2. Codes used to examine individual responses during peer-led conversations

Category	Individual Codes	Example
Relationship to overall conversation	Response to the original question	"I thought the main goal was to inform the visitor about public health…"
	Extending an idea	"Adding on to that, maybe there's a goal around us learning about public health…"
	Asking a new or clarifying question	"What do you mean by guidance?"
	Responding to the new or clarifying question	"You have the information, but don't know how to put it together…"
Emergent themes	Referencing the group	"Getting to know your ideas and how you see the world has been awesome too."
	Expressing vulnerability	"Coming to NYSCI, I was scared because I didn't want to come and join meetings."
	Critique of the program	"We need more guidance [from NYSCI staff] to be able to put our ideas into a form that works."

As shown in Table 3, the first peer-led conversation had the highest percentage of responses that were agreeing with or extending an idea. This suggests that these conversations were the "chattiest" with the most back-and-forth talk. Since this was the first group conversation and, knowing that the participants valued this time to talk and connect with one another, it could be that they had a lot that they wanted to say to one another. Additionally, this was the only session where the individual asking the questions had received training in facilitating conversations. Their input and prompting could have also contributed to the dynamic nature of the conversation.

Table 3. Relationship of a response to the overall conversation (% of total responses/conversation)

Conversation Number	Response to original question	Agreement with or extensions of an idea	New or clarifying question	Response to new or clarifying question
1	25%	47%	12%	20%
2	61%	29%	3%	3%
3	74%	11%	0%	0%

Table 4 suggests that peer-led conversations were prompting a large percentage of responses (29%-61%) that included elements of vulnerability. This was especially true for the second conversation, where more than half of responses included elements of vulnerability.

At the time the second conversation took place, fellows were focused on finalizing their idea for their exhibit experience and turning that into a prototype. This is something that was new for everyone and the conversation appears to have allowed them to share their struggles and questions with one another. The other cycles took place closer to the beginning and end of the program, when fellows were in the information-gathering or wrap-up stages of the program.

Table 4. Characterizing each response by emergent themes (% of total responses/conversation)

Conversation Number	References Group	Vulnerability/ Expressing a challenge or need	Critique of program
1	18%	29%	0%
2	13%	61%	3%
3	0%	30%	0%

Finally, program critiques were very limited across all conversations. However, the needs and challenges fellows shared—such as moments of confusion or areas where they wanted additional support—were often the most pressing issues on their minds and thus became the focus of their contributions. While these comments were not framed as formal critiques, they still provided valuable insight to inform program improvement.

Discussion

Overall, findings from this study suggest that peer-led conversations not only provided a comfortable place for youth to share their experiences, but also created rare opportunities for them to connect and reflect. Such opportunities were scarce in their busy schedules, partly due to logistical constraints (e.g., having to leave the building promptly at day's end and, in the spring, preparing for school exams) and were not part of the program's original design. These conversations also generated feedback that improved the fellows' experience in this and future programs.

Implications of these findings will be discussed in the remainder of this section.

Peer-led Conversations as a Means of Connection, Reflection, and Participation

According to Boud et al. (1985), reflection is the process through which "individuals engage to explore their experiences in order to lead to new understandings and appreciations" (p. 19). Reflection has been shown to have benefits in professional settings with the potential to improve practice by "identifying strengths and areas for improvement, changing processes, improving interpersonal and other skills, and increasing self-awareness and self-understanding" (Miller, 2020, p. 18). Reflection has also been incorporated in educational settings with students as a means of enhancing learning and engagement (Chakraborti & Pradhan, 2023).

Although fully involving youth in shaping and carrying out evaluation activities can be beneficial, it is not always appropriate for all contexts. Peer-led conversations can help provide some of the benefits of YPARE, such as self-reflection and connection (Odera, 2021), and also reduce some of the challenges of traditional forms of evaluation. Youth are talking with youth, lessening the youth/adult power dynamic. Having youth from the same cohort participating mutually in the conversations also reduces potential power imbalance in the group since, although one person is asking the questions, they are all reflecting on their shared experience. Additionally, the presence of the evaluator allows for the feedback to be shared with the project team while keeping participants' individual contributions confidential. Finally, this approach is less time-intensive to implement and requires less of an investment on the part of program staff, youth, and the evaluator.

Building on this, the approach presented in this study has the potential to be valuable for youths' learning and wellbeing in out-of-school settings, providing a method for evaluation that offers benefits to participants, gathers information for program improvement, and reduces some of the challenges associated

with YPARE. In this case, fellows' prior training and experience as peer facilitators likely contributed to the success of the approach, suggesting it may be particularly effective in programs that already emphasize youth leadership, peer engagement, and voice.

Challenges to the Use of Peer-led Reflections

The main challenge to the use of peer-led reflections was attendance, as not all fellows attended each session. As a result, some voices were missing from the conversation and also made it challenging to identify a facilitator in advance.

Youth voice was also somewhat absent from the design of the process and questions. Although evaluators attempted to seek their input in multiple ways, there were few responses unless there was a dedicated time to review and discuss the questions (e.g., during initial interviews, and at the end of their conversations).

Implications for the Field

This study of peer-led conversations has shed light on when and how this method may be an appropriate and beneficial data collection method as well as factors that may play a role in how participants respond to the conversation. The following recommendations are provided for those interested in implementing this method in their own work with young people:

1. **Consider whether the organization you are working with is ready and willing to implement peer-led conversations**. Implementing peer-led conversations requires that the organization you are working with is supportive of the method, understands the benefits to their program and youth, and sees it as an element that can and should be incorporated into the program structure. Supportive partners are vital to the success of this method because they need to provide the time and space for the conversations to happen, help with planning the conversations, express the importance of the conversations to youth participants, and be willing to use the feedback from youth to improve their experience in the program. Organizations and programs that value and forefront openness and continuous improvement would be especially strong partners, as this method aligns well with those values. Similarly, programs where youth participants already have experience with peer facilitation or leadership, as was the case with the fellows in this study, may find this approach particularly effective, since participants are better prepared to take on facilitation roles and engage meaningfully in the process.

2. **Build in time and money from the start to support peer-led conversations.** Evaluators and program staff should build this method into the program so that it is connected to and supporting the program experience. This takes planning from the very beginning of program development so that the time and money needed to support this method is available. This means building peer-led conversations into the program proposal as a means of supporting youth and understanding their experience.

3. **Be explicit with participants about the purpose of the conversations and how their responses will be used.** Evaluators should take time to explain that the conversations are a chance to talk and connect, while also providing feedback for the purpose of improving the program. If the project team makes changes in response to youth feedback, the team should mention that so that participants can see how the conversations are leading to direct change in the program.

In this study, for example, fellows' requests for greater clarity on institutional goals led staff to dedicate additional time to reviewing those goals. Similarly, their expressed need for more direction and support prompted the recruitment of additional fellows and more regular, explicit feedback from facilitators. These visible, in-the-moment adjustments not only addressed immediate needs but also demonstrated to fellows that their input had a tangible influence on the program.

4. **Include youth participants in the process as much as possible**. When possible, include youth throughout the stages of developing, implementing, and interpreting the peer-led conversations. This will look different in different contexts and for different participants. It could be as complex as directly involving youth in developing questions and interpreting findings or simply providing an opportunity for youth to review the questions and findings and providing a space for them to share their thoughts. Explaining to participants that the conversations are meant to support them, and engaging them in ways that allow them to contribute their thoughts about the evaluation, can help make the conversations more relevant and beneficial to them.

It is worth noting that the relatively small number of explicit program critiques does not suggest a reluctance to share negative feedback. Rather, the needs and challenges fellows voiced during the conversations reflected what was most salient to them at the time and were shared openly in the peer-led setting. For evaluators, this highlights the importance of allowing participants to guide the focus of feedback sessions toward the issues that matter most to them, even if these do not take the form of traditional critiques. Such input can still be highly actionable and directly inform program improvement.

Although peer-led conversations are not meant to fully replace other data collection methods, they do provide another way for youth voice to be heard. Youth will naturally speak differently to an evaluator versus their peers, so capturing these different voices from the same individuals can potentially add more depth to our understanding of their experiences. The results of this study indicate that peer-led conversations still shed light on youth experience and suggest actionable recommendations for program improvement, but they do so in a way that centers youth and their group connection.

Conclusion

In recent years, the evaluation field has made strides in being more inclusive of and responsive to young people, but we are still hampered by factors such as time, money, and the expectations of clients, funders, and even some of our colleagues and ourselves. We need to look critically at our evaluation designs and methods and focus on centering the participants.

Peer-led conversations are one way that youth voice can be heard in ways it typically is not. It provides the benefits of participants connecting and sharing with one another and also provides a look into their group experience. The open, vulnerable conversations and discussion on the group experience also led to better evaluation data and a deeper understanding of the experiences of the group and individuals.

The evaluation field needs to continue exploring methods such as peer-led conversations so that we can push our practice to move beyond extracting information. Valuing and acting on the issues that are most salient to participants—even when they are not framed as formal critiques—helps ensure that youth voices are not only heard but also shape meaningful and timely program improvements. Evaluators need to center the voices and lived experiences of evaluation participants and use methods that can offer benefits to the participants while still gathering high quality feedback and information to improve the programs that seek to be of value to them.

Corresponding Author: Jennifer Jocz, Educational Development Center
Email: jjocz@edc.org

References

Acree, J., & Chouinard, J.A. (2020). Exploring use and influence in culturally responsive approaches to evaluation: A review of the empirical literature. *American Journal of Evaluation*, 41 (2), 201-215. https://doi.org/10.1177/1098214019879505

Brewington, Q.L., & Hall, J.N. (2018). Givin' stakeholders the mic: Using hip-hop's evaluative voice as a contemporary evaluation approach. *American Journal of Evaluation*, 39 (3), 336-349. https://doi.org/10.1177/1098214018769765

Cargo, M., & Mercer, S. L. (2008). The value and challenges of participatory research: Strengthening its practice. *Annual Review of Public Health*, 29, 325–350. https://doi.org/10.1146/annurev.publhealth.29.091307.083824

Djohari, N., Simpson, L., & Simpson, K. (2020). Peer research with young people: Methodological and ethical insights. *International Journal of Social Research Methodology*, 23 (4), 443-456. https://doi.org/10.1080/13645579.2019.1678235

Hood, S., Hopson, R. K., & Kirkhart, M. W. (Eds.). (2015). *Hit hard: One hundred years of black baseball*. University of Nebraska Press.

Hook, J. N., Davis, D. E., Owen, J., Worthington, E. L., Jr., & Utsey, S. O. (2013). Cultural humility: Measuring openness to culturally diverse clients. *Journal of Counseling Psychology*, 60 (3), 353–366. https://doi.org/10.1037/a0032595

Kirk, S. (2007). Listening to young children's voices in research: Recognising children as participants, not just subjects. *Early Child Development and Care*, 177 (2), 173-181.

Knight, J. K., Wise, S. B., & Sieke, S. (2016). Group random call can positively affect student in-class clicker discussions. *CBE—Life Sciences Education, 15*(4), ar56. https://doi.org/10.1187/cbe.16-02-0109

Larson, R., Walker, K., & Pearce, N. (2005). A comparison of youth-led and adult-led discussions of adolescents' developmental experiences. *Journal of Youth and Adolescence*, 34 (3), 167-176.

Lesko, N. (2012). *Act your age!: A cultural construction of adolescence* (2nd ed.). Routledge.

Mason, M., & Hood, S. (2011). Youth-initiated research: A guide to the possibilities. *Journal of Community Psychology*, 39 (5), 629–642.

Mertens, D. M. (1999). Inclusive evaluation: A framework for holistic and ethical evaluation. *In International Handbook of Educational Evaluation* (pp. 133–155). Springer.

Murry, V. B. (2006). The impact of gender on the validity of self-reported delinquent acts. *Journal of Youth and Adolescence*, 35 (5), 831-840.

Rodriguez, S., & Acree, J. (2021). The importance of culturally responsive survey design in evaluation. *New Directions for Evaluation*, 2021 (172), 11-26.

Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking in conversation. *Language, 50*(4), 696–735. https://doi.org/10.2307/412243

Savage, J. (2007). *Teenage: The creation of youth culture*. Viking

Schelbe, L., Ononobi, F., & Duncan, E. (2015). Youth participation in qualitative research: Challenges and possibilities. *Qualitative Social Work*, 14 (4), 504-521. https://doi.org/10.1177/1473325014556792

SenGupta, S., Hopson, R., & Thompson-Robinson, M. (2004). Cultural competence in evaluation: An overview. *New Directions for Evaluation*, 102, 5-19. https://doi.org/10.1002/ev.112

Thomas, V.G., & Campbell, P.B. (2021). *Evaluation in today's world: Respecting diversity, improving quality, and promoting usability*. SAGE.

United Nations. (1989). *Convention on the Rights of the Child.*