A formative multi-method approach to evaluating training

Holly Hayes[a,*], Victoria Scott[b], Michelle Abraczinskas[c], Jonathan Scaccia[d], Soma Stout[d], Abraham Wandersman[c]

[a] University of South Carolina, 915 Green Street, Suite 506, Columbia, SC 29208, United States
[b] University of North Carolina at Charlotte, 9201 University City Blvd, Charlotte, NC 28223, United States
[c] University of South Carolina, Department of Psychology, 1512 Pendleton Street, Columbia, SC 29208, United States
[d] Institute for Healthcare Improvement, 20 University Road, Cambridge, MA 02138, United States

ABSTRACT

This article describes how we used a formative multi-method evaluation approach to gather real-time information about the processes of a complex, multi-day training with 24 community coalitions in the United States. The evaluation team used seven distinct, evaluation strategies to obtain evaluation data from the first Community Health Improvement Leadership Academy (CHILA) within a three-prong framework (inquiry, observation, and reflection). These methods included: comprehensive survey, rapid feedback form, learning wall, observational form, team debrief, social network analysis and critical moments reflection. The seven distinct methods allowed for both real time quality improvement during the CHILA and long term planning for the next CHILA. The methods also gave a comprehensive picture of the CHILA, which when synthesized allowed the evaluation team to assess the effectiveness of a training designed to tap into natural community strengths and accelerate health improvement. We hope that these formative evaluation methods can continue to be refined and used by others to evaluate training.

© 2016 Published by Elsevier Ltd.

1. Introduction

Training involves imparting specific knowledge, and skills to participants (Aswathappa, 2000) and is a commonly used means for professional development across American industries and sectors. In 2014, organizations in the United States spent $1229 on average per employee on training (ASTD, 2015). Despite the sizable and on-going organizational investment in training, studies on training outcomes show diminishing returns (Diamantidis & Chatzoglou, 2014; Grossman & Salas, 2011; Van Wijk, Jansen, & Lyles, 2008) and call into question ways that training outcomes can be improved.

Training effectiveness refers to the extent to which a training produces its intended results (Sitzman & Weinhardt, 2015). Typically, the effectiveness of a training program is measured at the end of the training program. For example, one of the most widely used models of training evaluation is Kirkpatrick's Four-Level Model which measures participant affective reaction, learning, behavioral change, and organizational impact post-training (Kirkpatrick, 1998). While post-training outcomes are critical to determining training effectiveness, in-training (formative) evaluations function to elevate training effectiveness. Formative evaluation methods provide real-time feedback about training programs. The timely feedback makes it possible for mid-course modifications to be made to the training. Formative evaluations can help improve a training program by illuminating gaps in the training program (e.g., instructional material and methods), unanticipated participant needs, and concerns about the training environment (Brown & Gerhardt, 2002). In-training evaluations are also known to enhance the transfer of training information back to the work place (Saks & Burke, 2012). Despite the important utility of formative evaluations for achieving desired training outcomes, there is a dearth of literature on formative evaluation methods. In fact, the evaluation phase is one of the most neglected aspects in training (McClelland, 1994) and is a major challenge for organizations (Aguinis & Kraiger, 2009; Twitchell, Holton, & Trott, 2000).

Over the last half-century, models for training have become increasingly more comprehensive as they have moved from focusing on outcomes to encompassing individual, contextual

* Corresponding author.
E-mail addresses: hayeshg@mailbox.sc.edu (H. Hayes), victoria.chien@gmail.com (V. Scott), michelle.abraczinskas@gmail.com (M. Abraczinskas), jonathan.p.scaccia@gmail.com (J. Scaccia), sstout@IHI.org (S. Stout), WANDERAH@mailbox.sc.edu (A. Wandersman).

and training interval factors that influence training outcomes (Cannon-Bowers, Salas, Tannenbaum & Mathieu., 1995; DeMatteo, Dobbins, & Lundby., 1997; Rowold, 2007; Scaduto, Lindsay & Chiaburu, 2008; Tai, 2006). Despite significant strides, however, the field continues to lack a model that fully captures a central quality of training, namely the *process* important to achieving effective training (Sitzmann & Weinhardt, 2015). Using a formative evaluation approach in which evaluation cycles occur during the training process can help address this gap by uncovering strengths and areas for improvement, which if addressed in real-time and between trainings, can facilitate desired training outcomes.

This article adds to the training literature by providing a formative evaluation multi-method approach to evaluating training. We describe our use of a three-pronged framework for evaluation and how our formative evaluation methodology was used to improve training outcomes in a large-scale, national health initiative. We conclude with recommendations for evaluating complex training programs. The major purpose of this article is to describe the methodology in detail to benefit the work of other community practitioners (e.g., trainers, organizational leaders, researchers, evaluators, etc). We do not focus on summative outcomes of the training due to the extensive literature that already exists on the topic and the specific gaps in reporting and publishing formative evaluations of training. Instead, we choose to demonstrate the use of formative evaluation to showcase how it can be helpful to practitioners and organizations.

2. Background: the community health improvement and leadership academy (CHILA) training – a component of the SCALE initiative

The Institute for Healthcare Improvement (IHI) is working with communities to understand how to promote and spread a culture of health through the development and implementation of a signature 100 Million Healthier Lives program called SCALE (Spreading Community Accelerators through Learning and Evaluation) (IHI, 2015). The SCALE Initiative, funded by the Robert Wood Johnson (RWJ) Foundation, is a unique 20 month intensive "learning and doing" program for 20 communities who are paired with an improvement coach and four mentor communities. The overall goal of the SCALE program is to support local leaders' success and to multiply their effectiveness in achieving their existing vision and goals. By participating in SCALE, communities, IHI and RWJ hope to learn what it takes to create the spread of useful approaches between communities in order to accelerate the pace of health improvement. A major component of the SCALE initiative is for communities to participate in four in-person training sessions over the 20 month period known as the Community Health Improvement and Leadership Academy (CHILA). The first CHILA took place between June 5 and June 9, 2015 in Wellesley, MA. An evaluation team (that includes the authors) is performing a formative evaluation of the overall SCALE program, including the CHILA.

3. Formative multi-method evaluation of CHILA

Since the CHILA evaluation is part of the larger SCALE formative evaluation, it was necessary to design an evaluation plan that was more comprehensive than just capturing training outcomes. The overall evaluation questions were: ***Were the support system components effective/sufficient for accelerating change? What worked and what didn't work?*** While these outcomes were important to document, we needed to gather information about CHILA experiences and processes in order to help improve the CHILA in real time. We also needed our evaluation to provide sufficient information to improve future iterations of CHILAs *and*

SCALE activities as a whole. We decided that using a multi-method approach for the evaluation would help us accomplish these diverse aims of real-time continuous quality improvement coupled with long-term improvement. Multi-method approaches use a variety of data sources and methods to collect different types of data (Bonoma, 1985; Morse, 2003); they provide a richer and broader understanding of what is being evaluated after triangulating results (Bonoma, 1985; Morse, 2003) to determine consistencies, discrepancies, and unique findings.

The overall evaluation approach was framed along three prongs (Inquiry, Observation, and Reflection); this approach has conceptual roots in Developmental Evaluation (Patton, 2006) and from strategies articulated by the Medical Research Council (Moore, Audrey, Barker, Bond, Bonell et al., 2015). Through *inquiry,* we actively sought information that answered our evaluation questions using survey and brief feedback forms. Through *observation,* we collected data on activities that naturally occurred during the CHILA, either by participating in selected events or reviewing documents that were produced by the SCALE implementation team. Finally, through *reflection*, we employed a structured narrative approach to obtain multiple stakeholders' input on how the CHILA worked and its critical drivers of success. This three-pronged framework for evaluation supported our interest in constructing robust narratives to capture the CHILA experience from multiple perspectives. In subsequent sections of this article, we discuss different evaluation methods incorporated to align with the three-pronged evaluation framework.

4. Participatory evaluation approach

We made a concerted effort to openly and directly communicate the spirit and values of the CHILA evaluation within a participatory evaluation framework to all stakeholders: the implementation team (those primarily responsible for designing and implementing SCALE activities to communities), the SCALE communities, and the funder. Our interactions with the SCALE implementation team consisted of participating in the CHILA design calls beforehand and seeking active input and feedback about our participatory evaluation design (Zukoski & Luluquisen, 2002). A collaborative process ensured that our evaluation processes would provide the implementation team with data they could use to improve the implementation of the first CHILA and future CHILAs, as well as capture elements that were important to them. For example, the implementation team conducted a pre-training needs assessment with communities to determine what would be most beneficial to accelerating their health improvement work. We used this information to structure specific evaluation activities to ensure that this content was adequately transferred.

When we solicited data from the communities (the process of *inquiry*), we communicated to participants the value of organizational learning and accountability (Fetterman 2015); our purpose was to use this data to both assess the progress of the communities, and the implementation of the CHILA so that methods to reach outcomes can be improved. Finally, the evaluation plan was reviewed by the funder (RWJ) to ensure that we were evaluating processes and outcomes that were priorities for them. By operating in an open and transparent manner with all key stakeholders, we facilitated a process that stressed the importance and utility of the data.

5. Methodology

For the first CHILA, the SCALE implementation team identified three training goals, which corresponded to the three primary drivers of success in the larger SCALE initiative. Each of these goals

was accompanied by a corresponding guiding question and objectives (Table 1). Seven distinct, mixed method evaluation strategies were used to obtain the process and outcome evaluation data from the first CHILA (Table 2). Each of the methods is described in detail below. See Table 3 for a logic model of the evaluation approach.

5.1. Evaluation technique 1: comprehensive questionnaire

An eight-page questionnaire was developed to examine the degree to which the three training goals included in Table 2 were met. The questionnaire was designed to be completed within 20 min, but participants could take as long as needed to complete the evaluation. Community members were asked to select which of the 24 communities they represented, but did not include any name identifiers on the survey.

This survey contained several sections. A retrospective pre-posttest was used within the eight-page survey to assess training effectiveness, which is believed to be more accurate in assessing perceived changes in participants over time than the common pre-test post-test methodology (Pratt, McGugan, Katzev, 2000; Goedhart & Hoogstraten, 1992), since it corrects for initial positive presentation bias. The survey contained ten statements rated on a Likert scale from 1 (strongly disagree) to 5 (strongly agree) that were pulled from the specific goals and objectives of CHILA. Participants were instructed to reflect upon their perspective prior and post CHILA when answering. The survey also included two Likert scale and two open ended questions each about the specific elements of CHILA (i.e. a matching process between community, the development of relationships with mentor communities, and visits to "bright spot" organizations) that occurred while at the first CHILA. Next, a 14 item instrument was used to assess the overall CHILA experience. Some of the statements that were rated on a scale of 1 (strongly disagree) to 5 (strongly agree) included: "the quality of the presenters and facilitators were excellent"; "working with peer groups enhanced my learning"; and "my overall experiences with CHILA was excellent". Lastly, open ended questions asked participants to describe the most valuable aspects of CHILA, what they most enjoyed, what they learned, how they planned to apply their learning, what they didn't learn and wished they had, and the least valuable aspects of CHILA.

5.2. Evaluation technique 2: rapid feedback form

To provide further information related to training goal one, a rapid feedback form (Fig. 1), was used immediately following speakers and workshops (modified from the University of Wisconsin-Extension, 2009). The sheets were printed on brightly colored paper and large stacks were left on each table during all $3\frac{1}{2}$ days. These forms asked for a quantitative assessment of the knowledge gained, value of the session, and practical application, along with a request for qualitative feedback on what participants intend to do as a result of the session. These forms were designed to take approximately two minutes to complete and were immediately collected and reviewed at the end of the session by the evaluation team. The qualitative information was summarized and presented to the implementation team during breaks in the CHILA.

5.3. Evaluation technique 3: learning wall: how are things shaping up?

As part of an inquiry and reflection activity, a Learning Wall was placed on the back of the main meeting room, which consisted of a 20′ by 7′ Technology of Participation© sticky wall that was displayed throughout CHILA. The purpose of the learning wall was to provide an open-ended opportunity where CHILA participants could post comments and thoughts about CHILA throughout the day, and read other participant comments. There were three specific prompts represented by shapes (Triangle, Circle, Square) that were used to frame participant's thinking. These questions can be found below in Table 4 and were modified from the University of Wisconsin-Extension (2009). Participants were asked to add post-it notes to the board throughout the day, and specific time was allotted for this reflection activity at the end of each day. All of the comments were collected, collated by category and reviewed by the Implementation and Evaluation Teams at a debrief session at the end of each day.

Table 1
CHILA Training Goals, Guiding Questions, and Objectives.

Training Goal	Guiding Question	Objectives
1) Develop leadership capability to transform and improve within and across communities at all levels.	To what extent did the first CHILA begin to develop participant capabilities as community leaders?	a) Begin to develop the inner conditions of leadership and partnership – identity, integrity, story of us, and story of now. b) Begin to unlock the capability in communities for leading sustained, joyful results-oriented improvement at the community level. c)Each community begins to develop a theory of change, driver diagram, logic model and a set of meaningful metrics to measure that change.
2) Create vibrant relationships and functional networks within and between communities that accelerate trust, learning, and achievement of a shared goal.	To what extent did CHILA1 begin to develop vibrant relationships among partners?	a) Help leaders and champions within and across communities begin to see themselves as a team. b) Start to develop smart peer-to-peer linkages as needed based on assets and needs and help communities roll up their community goals to a peer team goal. c) Start to create motivation, trust, and joy in the process of being together and creating change together. d) Start to create meaningful relationships between mentor communities, coaches and pacesetter peer group communities.
3) Promote bright spots between communities.	To what extent did the first CHILA begin to develop and identify bright spots (best practices and emerging best practices), and promote spread of bright spots between communities?	a) Begin to develop, identify, and connect community representatives to bright spots relevant to their work.

Table 2
CHILA Evaluation Methodologies.

Technique	Training Goal	Evaluation Category[2]	Type of Data[3]	Who completed	When Administered
Comprehensive Questionnaire	1, 2, 3	Inquiry	Process Outcome	Communities	Final Day
Rapid Feedback Form	1, 2, 3	Inquiry	Process Outcome	Communities	At the end of each session
Learning Wall: How are Things Shaping Up?	2	Inquiry	Process	Communities	End of each day
Observational Form	1, 2, 3	Inquiry	Process	Evaluation team	End of each day
Team Debrief	1, 2, 3	Observation	Process	Implementation team Evaluation team	End of each day
Social Network Analysis	2, 3	Inquiry	Outcome	Communities	Post-CHILA
Critical Moments Reflection	1,2,3	Reflection	Process Outcome	Implementation team	Post CHILA

Table 3
Logic Model for the Evaluation of the first CHILA.

Inputs	Evaluation Activities	Outputs	Outcomes
Evaluation team (8 individuals)	Comprehensive questionnaire	64 responses, represents 20 of the 24 committees	Actions taken by implementation team to improve CHILA 1 and CHILA 2 (See Table 5)
Weekly conference calls (~6 weeks prior to CHILA)	Rapid feedback form	1149 rapid feedback forms collected from 15 sessions	Participants were excited about the assortment of evaluation techniques, particularly that evaluation could be fun and provide just-in-time feedback. Many participants adopted the techniques for use with their communities.
Weekly calls with SCALE implementation team (for 6 months prior to CHILA)	Learning wall	Comments collected for each 3 categories for all 3 days	A more cohesive and integrated evaluation and implementation team (increased communication, trust, and understanding of roles)
SCALE implementation team and funder (RWJ) that was invested in the formative evaluation approach	Observations	Between 1–4 observation forms collected for the 15 sessions	
	Team debrief	3 team debriefs held after each day	
	Social network analysis	50 responses, represents all 24 committees	
	Critical moments reflection	10 participants by email, 4 participants in session, 59 critical moments generated, 2 critical comments chosen as most significant	

Fig. 1. Rapid Feedback form for all sessions.

5.4. Evaluation technique 4: observational forms

Observational forms were used during each session by evaluation team members. The forms included two sections: 1) What's working well in this sessions? 2) What changes or enhancements need to be made to this session? On the bottom of the form, space was provided for descriptive quotes from participants during closing debrief sessions and open discussions. Individuals kept all of their observational forms and they were then complied for analysis.

5.5. Evaluation technique 5: team debriefs

Immediately at the end of each day, an evaluation team member facilitated a brief (15–30 min) discussion with the SCALE implementation and evaluation team. The questions asked included: 1) What went well? 2) What could have gone better? 3) What surprised you/was unexpected? 4) What improvements need to be made? 5) Do you have other observations and comments? An evaluator had each of the questions printed on brightly colored paper and transitioned to the different questions based on the conversation emerging with the group. An evaluation team member documented all the comments made each day related to the five questions.

5.6. Evaluation technique 6: relationship survey

In an effort to further examine the extent that relationships were formed at CHILA (Training Goal 2, see Table 2) and bright spots shared (Training Goal 3, see Table 2), a relationship survey was distributed to all participants two weeks following CHILA. The brief survey administered via SurveyMonkey asked participants to select the names of communities from a roster that he or she connected with during CHILA. A connection was defined as sitting at the same table, sharing a table at a meal, having conversations with or eating together. In addition, participants were asked to report the names of communities that shared information with them that they planned to use (e.g. project descriptions, ideas, tools, data) within the next three months. An open text box was

Table 4
Questions on the 'How are Things Shaping up?' Learning Wall.

Shape	Question Headings on Learning Wall
Triangle	What are the 3 key points that really hit home? What insights emerged for you?
Circle	What is still circling in your head? What is still unclear for you today?
Square	What is squared away with you that you will apply and use within the next 3 months?

Table 5
How Formative Data was used by the Implementation Team.

Data from the first CHILA formative Evaluation (i.e. "You said,")	Actions taken by Implementation Team, (i.e. We Did,") during the first CHILA (Real Time Changes)	Actions taken by Implementation Team, (i.e. We Did,") during the second CHILA (long term changes)
The days were too long and we needed time to absorb information	Energizers added throughout the day, including ones contributed by the SCALE communities.	Agenda for CHILA2 had fewer long days, more time for reflection and digestion built in. There was also additional time for informal relationship building
Decrease information overload.	Orientation times added to the beginning of each day.	More breakout sessions were offered to customize learning and decrease information overload
People are at different stages of learning	CHILA1 served as basic foundational grounding.	We built a "refresher" session into CHILA2 for less experienced Local Improvement Advisers; created breakouts to meet people where they were at.
Plan additional time for self care (e.g. exercise, nap, check in with work/home) during future CHILAs	N/A	Hour and a half break each afternoon was added to create digestion time to give time for self-care
Better preparation and logistics for the site visits	N/A	Much more advanced planning with Bright Spot sites; much more knowledge provided about SCALE
Discuss plan for CHILA and community roles at beginning, and explain rationale behind all activities clearly.	Introduced "balcony" view was provided every day, hosted by consistent MCs. On the last day, S. Stout presented the overall journey.	A "balcony" view was provided every day, hosted by consistent MCs. A "balcony" view helps participants reflect on the dynamics they observe within their own process by taking on a perspective as if one were looking down on a stage from the balcony.
Relationship building was highlighted by communities as an important CHILA component.	Activities created with a focus on addressing equity.	More dedicated time for peer community teams at CHILA2
Communities want additional time at CHILA to build on and establish new relationships between and within communities.	Sometime built into CHILA1 for this.	More time as a home team to work on health improvement projects, more opportunities to invite team at home to call in at certain times

also available for participants to describe specific take-aways that they plan to apply before the next learning collaborative.

5.7. Evaluation technique 7: critical moments reflection

Approximately three weeks after the conclusion of CHILA, a Critical Moments Reflection session was held virtually over a web-based conference platform with key members of the Implementation Team and SCALE community partners who were involved in developing and implementing the CHILA curricula. The Critical Moments Reflection methodology attempts to raise the group consciousness about incidents, moments, and reactions that were perceived as being especially significant to the current status of the project (McDowell, Nagel, Williams, & Canepa, 2005). The rationale

for using this process was to surface important events that might not have been otherwise captured through the other evaluation methods.

To stimulate reflection, the following question was posed to stakeholders individually beforehand: *What are the most important learnings about the onboarding process [the preparation phase leading into the first CHILA] and the first CHILA from my perspective as someone closely involved in developing and implementing this phase of SCALE?* Stakeholders were then asked to submit five moments that they perceived as answering this question. The evaluation team then grouped these moments by theme and displayed them as specific markers of the CHILA (e.g. opening plenary).

During the critical moments session, a member of the evaluation team facilitated a conversation allowing each

participant to voice their critical moments that were rated either as positive or negative. Each stakeholder presented their critical moments and discussed how they saw these as influencing the processes and outcomes of the first CHILA. Following the review of the data submitted, a voting process was held to determine which moments were "most critical," and worthy of more in-depth discussion and processing by the group. Specific emphasis was placed on how lessons from these moments can be used to inform planning for the next CHILA. The number of critical moments that were discussed was constrained by the amount of time that was dedicated to this session (1.5 h). Other projects may discuss more moments, if appropriate.

The value of combining mixed evaluation methods allows one to paint a more comprehensive evaluation picture and provide both rapid feedback as well as information for long term planning for the initiative. One of the methods alone is not designed or able to capture the breadth and complexity of the training. Multiple methods allow one to validate observations and to ensure that one voice or perspective is not dominating the feedback.

6. Results

6.1. Evaluation technique 1: comprehensive questionnaire

Our results describe the outputs from the evaluation activities (refer to Table 1). The final CHILA evaluation survey was administered via paper form before the lunch break on the final day, June 9th. Sixty-four responses were gathered, representing 20 of the 24 communities present at the first CHILA. Participants took 15 min to 45 min or longer to complete the survey. Each item in the pre- to post-assessment was statistically significant, indicating that participants were able to acquire additional knowledge and skills over the course of CHILA. In addition, participants rated their overall experiences with the CHILA both quantitatively and qualitatively. Tables were created in the evaluation report showing the average ratings of each item along with qualitative comments grouped by theme. Themes from the seven open-ended qualitative items at the end of the survey were summarized into a two page narrative with a descriptive visual.

6.2. Evaluation technique 2: rapid feedback form

The Rapid Feedback Forms were completed after specific sessions designated by the Principal Investigator. Due to time restraints, certain sessions did not have rapid feedback forms completed. The evaluation team collected 1149 rapid feedback forms for 15 of the sessions held over the three days. The average scores and standard deviations for each of the three factors (knowledge gained, value of session and practical application) were included in the evaluation report. Qualitative data placed on the back of the Rapid Feedback Forms were synthesized by session and provided in the full evaluation report.

6.3. Evaluation technique 3: learning wall: how are things shaping up?

The implementation team reviewed the post-its from the Learning Wall and reviewed them during the daily debrief. Themes that emerged for the Triangle (i.e. learning from mistakes, shifting focus of health improvement work), Square (i.e. learning about the improvement process, fostering an environment of joy) and Circle (i.e. large scope of health improvement work, challenge of building collaborative teams) were synthesized and included in the evaluation report. The Learning Wall was used each day with a great degree of variability. We had the most participation for the activity on the first day (over 70 post-its) and the participation dropped to less than 10 post-its on the final day.

6.4. Evaluation technique 4: observational forms

Observation forms were collected on each of the 15 sessions, with one to four observers for each session. Some of the evaluation team members typed this for the sessions they were responsible for and others wrote their observations on paper for each of their sessions. Each evaluation team member was asked to pay careful attention to the number of participants in each session, laughter, questions asked, and interaction among community members. Evaluation team members also used their personal cell phones to capture some video clips from the energizers held between sessions (brief physical activities/stretches designed to reengage participants). Photos were also taken of participants and speakers and some were included in the evaluation report.

6.5. Evaluation technique 5: team debriefs

De-brief meetings were held with the implementation and evaluation team after each day. These sessions were recorded by one of the evaluation team members and then used to supplement data from the observation and rapid feedback synthesis in the report.

6.6. Evaluation technique 6: relationship survey

The relationship survey was distributed through Survey Monkey. There were a total of 50 respondents representing all communities, with a mean of 2.08 respondents per community (range 1–5). Communities were identified that had the most and least new relationships built at CHILA. Communities also were identified whom shared practical information at CHILA that was then put to immediate use in the home communities. The evaluation team decided that only the communities selected most often were shared in the report by percentage and not communities selected the least often, since the evaluation report would be seen by the implementation team and community members. The evaluation team did not want to make any community "feel bad" for being selected the least. A visual was created to illustrate the informal connections using GIS with the map of the United States showing connections between the 24 communities. In addition, specific tools being shared by communities were collated and presented in the evaluation report.

6.7. Evaluation technique 7: critical moments reflection

The critical moments reflection methodology had two phases. A total of ten implementation team members submitted critical moments prior to the session. Although participants were asked to submit five moments each, the number of submissions ranged from 1 to 11, with the average number of moments equaling 5.9. Of the ten people who submitted moments beforehand, four attended a session facilitated by the evaluation team. The two topics that were voted upon as "most critical" (*the intentionality of the design process* and *the peer matching process*) were discussed in detail, including the implications of these critical moments for future learning collaboratives. The session was digitally recorded. We created three visuals that displayed the frequency that stakeholders identified moments grouped by theme, and mapped to specific time points during CHILA. These visuals were included in the evaluation report.

7. Real time changes & long-term planning

Implementing the evaluation plan *during* CHILA took the coordinated effort of four members of the evaluation team. The level of collaboration with the implementation team was

instrumental in a number of ways. First, during the CHILA, the implementation team helped emphasize the spirit and importance of the evaluation by voicing their support for all evaluation activities. For example, evaluation was not communicated as a burden or add −on, but rather as an instrumental method of co-learning and improvement. Second, the implementation team helped to assemble and distribute materials ahead of time. Third, they supported the evaluation by providing time during sessions for the completion of evaluation activities.

The comprehensive survey allowed the evaluation team to gain rich data related to the three training goals (Table 1). Some participants voiced that they needed more time to complete the survey and took the survey with them to lunch (some never returned). More time should have been given to the participants to complete the survey, which took an average of 30 min to complete. One participant noted, "This is important, and we need to give this time to really think about." Participants were also extremely fatigued and overwhelmed the final day, noted one of the community members. The costs and benefits should be considered of having a captive audience and allowing data to be immediately gathered versus giving participants time to complete the survey at their convenience within a week of returning home. In future training sessions, we will have both a survey near the end of the CHILA session and email community members one week following the training; this will provide additional time to reflect whether they have additional insights or further recommendations for future training sessions.

The Rapid Feedback Form was very successful in getting quick feedback on the sessions and took an average of less than two minutes to complete. After the first session, it became apparent that the Principal Investigator and the Implementation Team wanted to hear feedback to guide future sessions. Following this realization, an evaluation team member was assigned to quickly review the data after each round of form collection and provide a one-page handwritten summary to the Principal Investigator with overall feedback of themes emerging from the qualitative data, within a half hour after the session was over. This revised practice proved to be very helpful to the Implementation Team, who quickly made some changes based on feedback (e.g., too many instructions being given by different people, honoring silence, need for breaks, wanting more time for discussion). In future iterations of the form, we may modify the center statement and add a question that reads "Any additional feedback to share with the planners?" There was not a place for additional comments, and often participants wrote comments related to noise level, length of presentation, and need for additional information on the back of the form.

The implementation team gained insights and reviewed feedback from the Learning Wall. Due to time constraints and the increased level of individuals' activities, the reflection time was only given on day one. Several real time changes were made on day one based on input from the learning wall related to water being placed on tables, more protein at snacks and the need for more breaks. For the remaining days, the learning wall was mentioned by implementation team members in passing with the morning debrief. Community members were also experiencing some fatigue from participating in many interactive sessions and being asked to reflect during short breaks before starting another workshop. Following the training, we were informed that at least three communities have used the learning wall with their own communities. The implementation team felt that this evaluation activity made the evaluation team more approachable.

Over 10 pages of the 61-page final evaluation report of the first CHILA were dedicated to an observation synthesis of each session combined with a synthesis of the data from the rapid feedback form. Implementation team members and speakers found this level of detail extremely important in "telling the story" and also to

give specific feedback to speakers, who will continue to facilitate and lead sessions for upcoming training sessions. For future training sessions, it is important to detail what evaluation team members should be documenting (interaction, numbers of participants, noise level, etc.) and to create a detailed schedule ahead of time of which sessions will be observed by whom. Two of the evaluation team members experienced some fatigue observing three or more sessions back to back without a break.

The daily team debriefs strengthened the relationship between the evaluation and implementation teams and created a deeper sense of trust and understanding of each other's roles. As a result of the debriefs, real time changes included having a single person share announcements, providing more explanation of activities, and giving more time for participants to share their community expertise. It is important to make sure that every team member is aware of the value of the team debriefs from a continuous quality improvement and relationship- building perspective ahead of time. It was clear that team members were expected to participate in the debriefs to help build a positive learning experience and environment for the community participants.

The findings from the relationship survey proved to be problematic when trying to draw conclusions from this initial social network analysis survey. The results of the survey were only intended to show broad connections that were formed during the learning collaborative. For future iterations, the evaluation team will limit the survey to one community member (i.e., the coalition leader), since the participation of future training sessions events will vary.

The richness of information and discussion using the critical moments methodology is dependent on having a sufficient number of participants. Unfortunately, the critical moments reflection took place on a Thursday before a holiday weekend, and the majority of the team could not participate in the reflection activity. In the future, we will only host critical moment reflections when we have at least eight participants and will also send out calendar invites a month ahead of time. Ground rules are also important to promote sharing of critical moments in a safe and protected space, which is especially beneficial if one moment conflicts with another individuals' critical moment (e.g., one moment could be positive for one and negative for another). The key for this reflection is to have a trained facilitator guide the discussion grounded in the central agreed upon question.

Overwhelmingly, the first CHILA was perceived as a positive experience by the SCALE communities. Feedback was nearly universal in recognizing the value of forming between- community relationships and becoming part of a larger social movement to address meaningful change in health-related outcomes. A 61-page evaluation report was developed for the SCALE key stakeholders (implementation team, RWJF, and the SCALE communities) which included detailed narrative, data tables, charts, figures and an Appendix with qualitative data responses grouped by themes. This report was shared by email and also posted on the collaborative web platform used by communities. Furthermore, the implementation team developed a set of "You said; We did," slides to share with communities (see Table 4 for a summary of these slides.) The purpose of this activity was to communicate to the communities that their feedback was read and directly used to inform changes in how SCALE and CHILAs would be implemented in the future. These slides were presented in the opening session of the second CHILA.

8. Conclusions from the multi-method approach

Critical for the evaluation of the training sessions was using a multi-method approach and communicating regularly with the implementation team. It was very important for the evaluation team to be modeling collaborative, engaged behavior at the CHILA.

Evaluation team members actively participated in the sessions just like the community members, which increased community members' comfort level with the evaluation team. This also allowed the team to gain significant insight into the intensity and dosage of the training. In addition, the implementation team and the evaluation team worked closely together all $3^1/_2$ days providing feedback to one another in a very timely manner which created a culture of continuous quality improvement. One implementation team member noted that "there were no surprises" from the evaluation report, since the evaluation team had done a good job sharing feedback with the team in a very timely manner. Improvement of the training event between sessions was a by-product of the real-time feedback to the implementation team throughout the $3^1/_2$ days. The relationships formed between the evaluation and implementation team at the CHILA have strengthened over time and greatly assisted with future formative evaluation work.

The methodology used for synthesizing the evaluation data was critical in ensuring high utility of the final report. With seven distinct evaluation methods, it is easy to drown in synthesizing the vast amount of data gathered in different ways and from different types of participants (community members, implementation team members). Initially, our evaluation team moved quickly into data entry, compiling and synthesizing. We quickly learned that while the multi-method evaluation approach provides a broad range of data related to the training, the importance of understanding how all of the data will be used becomes paramount. During an evaluation team meeting, it was determined that we had to work "right to left" instead of "left to right." "Right to left" thinking helps a group think about what the end product needs to look like and the major outcomes and then work backwards to achieve the outcomes. "Left to right" thinking, in contrast, focuses on where we currently are and what next steps need to happen. While both thinking patterns are needed, it was important for our group to agree on what the evaluation report would look like and its intended audience. The evaluation team agreed that the initial evaluation draft would be written to all audiences (funder, implementation team, and community participants) and different executive summaries and PowerPoint slide decks would be created for different audiences.

It is important to take time to reflect on all of the results gathered from the different methods to craft thoughtful recommendations following a training. In addition, the evaluation team believed it was critical to submit a product to the Principal Investigator and Implementation team within four weeks of the training, since plans were being made for the second CHILA training to be held three months after the first CHILA training. In the evaluation report, recommendations were made in three categories: 1) Results, 2) Process, and 3) Relationships; this model from Interaction Associates was used throughout the development of the learning collaborative. By reporting the recommendations in a framework that the implementation team uses regularly, the recommendations were easier to understand and apply. After the final evaluation report was shared with communities, the Principal Investigator highlighted specific changes that were being made at the upcoming training based on their evaluation feedback.

One unintended outcome that emerged after CHILA is how communities are learning about evaluation through the experience of being involved in an evaluation. We have been very explicit with all stakeholders about the specific rationale behind our evaluation methods, the spirit and values of our evaluation, the results of the data that we collect, and how we intend to use the data. Although the CHILA did not explicitly have a goal about evaluation capacity building, communities approached us both during and after the CHILA with questions about how they could use these methods in their own settings. We now plan to track evaluation-capacity building over the course of SCALE as a consequence of experiencing evaluation. Evaluation capacity will be tracked though interviews with case study communities with a specific emphasis on evaluation training transfer between the four CHILAs.

While we were fortunate to have a deeply collaborative relationship between implementers and evaluators in SCALE, the techniques identified in this article for rapid cycle formative evaluation are relevant for practitioners anywhere who are trying to implement a new curriculum or new educational methods. Both the thematic presentation of feedback (often within 10 min of the end of a session) and the detailed, rich feedback of participant experience have been invaluable in creating a learning process that has enriched both learners and educators engaged in SCALE. While we do not report on training transfer in this article, we do address it in our larger formative evaluation plan. We recognize that the transfer of knowledge following or between training sessions is an important area to assess and ultimately demonstrate the training's long term effectiveness.

The quick continuous quality feedback loop implemented at CHILA provided mid-course corrections that may improve long term cognitive and behavioral training outcomes (Salas & Stagl, 2009) because the corrections were geared toward addressing the attendees' learning needs. We strongly encourage training evaluators to consider adding rapid cycle continuous quality improvement methods to their evaluation plans. Suggesting training improvements to stakeholders while the process is still occurring can build collaborative relationships and increase attendee satisfaction in training when they see that their feedback has been heard and used.

Consistent with current training literature (Salas, Tannenbaum, Kraiger, & Smith-Jentsch, 2012), the implementation team, IHI, and the RWJ all consider training integrated into the project, rather than a separate component. The use of multiple methods is especially fruitful in this type of project because it captures the full scope of participants' experiences. The seven distinct methods allowed for both real time quality improvement and long term planning for the next CHILA. The methods also gave a comprehensive picture of the CHILA, which when synthesized allowed the evaluation team to assess the effectiveness of a training designed to tap into natural community strengths and accelerate health improvement. We hope that these methods can continue to be refined and used by others to evaluate training.

References

Aguinis, H., & Kraiger, K. (2009). Benefits of training and development for individuals and teams, organizations, and society. *Annual Review of Psychology, 60*, 451–474.

American Society for Training & Development (2015). *2015 State of the industry.* Alexandria, VA: ASTD Research Department.

Aswathappa, K. (2000). *Human resource and personnel management.* New Delhi: Tata Mcgraw-Hill Publishing Company Limited p. 189.

Bonoma, T. (1985). Case research in marketing: Opportunities, problems and a process. *Journal of Marketing Research, 2*(2), 199–208.

Brown, K. G., & Gerhardt, M. W. (2002). Formative evaluation: An integrative practice model and case study. *Personnel Psychology, 55*(4), 951–983.

Cannon-Bowers, J. A., Salas, E., Tannenbaum, S. I., & Mathieu, J. E. (1995). Toward theoretically-based principles of trainee effectiveness: A model and initial empirical investigation. *Military Psychology, 7*, 141–164.

DeMatteo, J. S., Dobbins, G. H., & Lundby, K. M. (1997). The effects of accountability on training effectiveness. *Training Research Journal, 3*, 39–57.

Diamantidis, A. D., & Chatzoglou, P. D. (2014). Employee post-training behaviour and performance: Evaluating the results of the training process. *International Journal of Training and Development, 18*(3), 149–170.

Fetterman, D. M. (2015). Empowerment evaluation: Theories, principles, concepts, and steps, 2nd ed. In D. M. Fetterman, S. J. Kaftarian, & A. Wandersman (Eds.), *Empowerment evaluation:* (20–42. pp. Thousand Oaks, CA: Sage.

Goedhart, H., & Hoogstraten, J. (1992). The retrospective pretest and the role of the pretest information in evaluation studies. *Psychological Reports, 70*, 699–704.

Grossman, R., & Salas, E. (2011). The transfer of training: What really matters. *International Journal of Training and Development, 15*(2), 103–120.

Institute for Healthcare Improvement (2015). *IHI announces final selection of communities for health-focused SCALE initiative.* Weblink: http://www.ihi.org/about/news/Documents/IHIPressRelease_IHIAnnouncesSCALEAwardees_Apr15.pdf.

Kirkpatrick, D. L. (1998). *Evaluating training programs: The four levels*, 2nd ed. San Francisco: Berrett-Koehler.

McClelland, S. (1994). A model for designing objective-oriented training evaluations. *Industrial and Commercial Training, 26*(1), 3–9.

McDowell, C. L., Nagel, A., Williams, S. M., & Canepa, C. (2005). Building knowledge from the practice of local communities. *Knowledge Management for Development, 1*(3), 30–40.

Moore, G. F., Audrey, S., Barker, M., Bond, L., Bonell, C., Hardeman, W., . . . Baird, J. (2015). Process evaluation of complex interventions: Medical Research Council guidance. *British Medial Journal, 350*, h1258.

Morse, J. (2003). Principles of mixed methods and multimethod research design. In A. Tashakkori, & C. Teddie (Eds.), *Handbook of mixed methods in social and behavioral research* (pp. 189–208)..

Patton, M. Q. (2006). Evaluation for the way we work. *Nonprofit Quarterly, 13*(1), 28–33.

Pratt, C. C., McGuigan, W. M., & Katzev, A. R. (2000). Measuring program outcomes: Using retrospective pretest methodology. *American Journal of Evaluation, 21*(3), 341–349.

Rowold, J. (2007). Individual influences on knowledge acquisition in a call center training context in Germany. *International Journal of Training and Development, 11*(1), 21–34.

Saks, A. M., & Burke, L. A. (2012). An investigation into the relationship between training evaluation and the transfer of training. *International Journal of Training and Development, 16*(2), 118–127.

Salas, E., & Stagl, K. C. (2009). Design training systematically and follow the science of training, In E. Locke (Ed.), *Handbook of principles of organizational behavior: Indispensible knowledge for evidence-based management* (pp. 59–84).2nd ed. Chichester, UK: Wiley.

Salas, E., Tannenbaum, S. I., Kraiger, K., & Smith-Jentsch, K. A. (2012). The science of training and development in organizations: What matters in practice. *Psychological Science in the Public Interest, 13*(2), 74–101.

Scaduto, A., Lindsay, D., & Chiaburu, D. S. (2008). Leader influences on training effectiveness: Motivation and outcome expectation process. *International Journal of Training & Development, 12*(3), 158–170.

Sitzmann, T., & Weinhardt, J. M. (2015). Training engagement theory: A multilevel perspective on the wffectiveness of work-related training. *Journal of Management* 0149206315574596.

Tai, W. (2006). Effects of training framing, general self-efficacy and training motivation on trainees' training effectiveness. *Personnel Review, 35*(1), 51–65.

Twitchell, S., Holton, E. F., & Trott, J. W. (2000). Technical training evaluation practices in the United States. *Performance Improvement Quarterly, 13*(3), 84–109.

University of Wisconsin-Extension, Cooperative Extension, Program Development and Evaluation. (2009). Creative or alternative methods for collecting data. Web link: www.uwex.edu.

Van Wijk, R., Jansen, J. J., & Lyles, M. A. (2008). Inter-and intra-organizational knowledge transfer: A meta-analytic review and assessment of its antecedents and consequences. *Journal of Management Studies, 45*(4), 830–853.

Zukoski, A., & Luluquisen, M. (2002). Participatory Evaluation: What is it? Why do it? What are the challenges?. *Policy & Practice 5*, . http://depts.washington.edu/ccph/pdf_files/Evaluation.pdf.