Modifying the Process Evaluation Framework for Utilization in Extension and Social Sciences

Stacey F. Stearns, University of Connecticut – Storrs
Cristinia Connolly, University of Connecticut
Jeantyl Norze, University of Arkansas

Abstract

Medical fields use process evaluations to measure intervention fidelity as they are often complex with multiple interacting components. While Cooperative Extension programs are similarly structured, process evaluations have not been widely used in Extension or other social sciences but could be beneficial in improving programming using a modified framework. We modified Carroll et al.'s framework to conduct a process evaluation of an Extension program instead of in a medical field. This paper outlines the implementation fidelity using ex-post, semi-structured interviews conducted based on the process evaluation framework. The interviews helped complete the process evaluation and address data gaps.

Keywords: Process Evaluation, Program Evaluation, Cooperative Extension, New Technologies in Agricultural Extension, Game-based Learning, Team Development, Program Fidelity, Program Adaptation

Introduction

High-functioning teams focus on project outcomes, and team members' experience, including leadership on other teams, correlates to the success of a new team (Klug & Bagrow, 2016). While we understand that high-functioning teams need focus, leadership, and experience; less is understood in Extension about creating frameworks and processes that help them thrive. Duerden and Witt (2012) discussed how assessing program implementation is equally important to program impact, while Lamm and Israel (2013) outline that understanding the process can improve activities but is different from accountability evaluation. Therefore, the process evaluation is useful for the implementation team and as a decision-making tool, however, teams must see a use for the data to conduct a process evaluation (Lamm & Israel, 2013). Some researchers are using process evaluations, including on a social marketing campaign in Australia (Dietrich et al., 2019) and on a Mississippi Extension social marketing campaign (Downey et al., 2021). However, these examples were for the program and not the team itself.

Extension strives for successful projects with good team dynamics and outputs, and process evaluations could help teams achieve greater community impacts or change. Process evaluations determine how closely an implementation adheres to the original plan, and how much the process is responsible for program results and impact (Carroll et al., 2007). The medical field uses process evaluation frameworks systematically to address complex programs and improve intervention adherence (Hulscher et al., 2003; Linnan & Steckler, 2002; Moore et al., 2015). Social sciences are not using process evaluation frameworks as readily (Downey et al., 2021; Duerden & Witt, 2012; Lamm & Israel, 2013); yet an adapted version could help teams improve.

Process evaluation frameworks in the medical field carefully delineate the intervention and measure ten variables to examine implementation success rates: content, exposure/dose delivered, dose received, reach, quality of delivery, participant responsiveness, complexity, facilitation strategy, recruitment, and context (Carroll et al., 2007; Hasson, 2010; Hulscher et al., 2003; Lewis et al., 2020; Linnan & Steckler, 2002; Moore et al., 2015). Medical professionals using these variables in a process evaluation can determine why an intervention worked, or where it improved (Carroll et al., 2007; Hasson, 2010; Moore et al., 2015).

Determining the elements that are essential for success is key when trying to judge implementation fidelity while allowing for adaptability to local conditions. In this paper, we strove to use Carroll et al.'s (2007) process evaluation framework to answer the following research questions:

1. What can we discover and improve for greater impact through process evaluation?
2. What led to improved team dynamics?
3. Did our project activities successfully accelerate our development of a learning game, as well as external funding to support our Extension programs?
4. What could have further improved the team dynamics?

Program Description

The New Technologies in Agricultural Extension (NTAE) is a project accelerator for Extension teams, helping them work together more cohesively in achieving an outcome deliverable and to be competitive in future grant funding processes. Extension professionals secure a one-year fellowship, including access to a suite of wraparound services that provide marketing, leadership, digital engagement, publishing, partnership, and project mentors. Two NTAE “catalysts,” who are

retired Extension directors, serve as team mentors. Their responsibilities include developing leadership capacity in the team members and accelerating the project overall. Their role also provides continuity and collaboration among the NTAE staff's other wraparound services.

Our planned deliverable for the University of Connecticut (UConn) Extension NTAE project was a consumer game providing education about food marketing labels, as well as two field guides concerning food marketing labels and game development activities. The role of the catalysts was to support us as we developed the game and improve our team's performance, while the wraparound services served as consultants to help us build our skillset; the main purpose of the NTAE program is to improve team dynamics and capacity. We could then apply these skills, such as marketing or leadership, to the development and deployment of our game.

Members of the team met monthly with the NTAE staff to check-in, discuss new achievements, and detail the next steps. Our team also held monthly meetings on our own to facilitate the work. Sub-committees associated with specific roles on the team met more frequently and reported back on various components of the project at team meetings. There were synergies between all the wraparound services that fostered an environment for team growth and project acceleration. We partnered with the Learning Games Laboratory at New Mexico State University (NMSU) for game development, including a game jam. The lab is part of the NMSU Extension system and a leader in game-based learning. A key element of the design process entailed collaborating with NMSU designers and developers to create the game prototype. Our team developed a backward plan (or plan of work) at the beginning of the project year to structure the goals, objectives, activities, outcomes, and evaluation measures prior to commencing work (Fink Chorzempa, Smith, & Sileo, 2018; Wiggins & McTighe, 2005). A key measure of implementation fidelity is the degree to which the initial plan is followed; the backward plan aided with this.

Methodology

Population, Sample, and Instrumentation

This is a cross-sectional study that examines the quality of the implementation of the NTAE project through the lens of a process evaluation framework developed by Carroll et al. (2007) and expanded upon by Hasson (2010) and Moore et al. (2015). The study's target population was the project stakeholders including project staff, funders, participants, and partners**.** The researchers used the census or complete enumeration method for this study since the overall population (N = 21) was accessible.

We adapted this medical process evaluation framework, adjusting the variable definitions to apply to an Extension team, to measure the success of our project team and program implementation, and provide insights on replicating this success. While the definition of each component is the same, we aligned them with different parts of the program. For example, dose delivered in medical process evaluations may refer to the patient medications, whereas our definition is the meetings and training available to the team. To begin, we analyzed the existing backwards plan and meeting notes. We then developed interview questions and conducted ex-post facto (after the fact), semi-structured interviews to obtain additional information.

Our final interview questionnaire consisted of two sections. The first section included questions about interviewees' professional demographics. The second section included measures of the selected process evaluation framework as it relates to the delivery or implementation of the NTAE project. Appendix 1 details the interview questions and corresponding measures. The interview questionnaire was designed and developed by a research team consisting of a subset of

the NTAE team including the team leader, a team member, and two Evaluation Specialists, one of whom was also a team member and retired during the process evaluation. Once the research team agreed on the framing and the content of the questionnaire, it received Institutional Review Board approval and potential study participants were contacted.

Data Collection and Analysis

Emails invited the project stakeholders (N = 21) to participate in this study via interviews. Overall, 10 (47.62%) individuals participated in a semi-structured interview, which lasted forty-five minutes on average. The interviews were conducted via Zoom video and recorded for transcription and analysis. A student worker manually transcribed the ten video files using verbatim transcription and shared with the research team for analysis. The research team then developed a team protocol and used preset codes and code descriptions to analyze the interview data. They also developed a codebook using the selected process evaluation framework (Carroll et al., 2007; Hasson, 2010; Moore et al., 2015).

Each member of this same research team conducted individual coding (independently) using NVIVO software. We then convened for focused coding to align our individual coding results. Next each team member was responsible for analyzing and distilling the results of a subset of codes, at which point the research team convened again to discuss final results. Involving multiple researchers in a qualitative study provides a multifaceted understanding of the data and minimizes potential biases in the study findings (Tracy, 2019). The application of codes (Appendix 2) from the selected process evaluation framework to the study dataset was done through multiple iterations; provisional coding helped the team collectively determine how qualitative data aligned with each code. Codes were grouped together into clusters around similar and interrelated concepts or elements of the selected process evaluation framework.

Findings

Our process evaluation has eight components, which we analyzed individually and then as a whole. *Recruitment* concerns the plan to attract participants and any participation barriers, though in this case, it was primarily an ad hoc process that relied on networking. Of seven potential team members approached due to a perceived overlap with their expertise, all agreed to participate. We expect that personalized invitations contributed to this high level and the fact that participants were recruited based on pre-existing knowledge and interest in the topic.

Reach refers to the percentage of the target audience that participates in the intervention, including consideration of sub-groups. In the context of an Extension team, this could be the percentage of targeted personnel that joined or collaborated with the team and the degree to which different competencies or subject areas were represented. The team leader was intentional about creating an interdisciplinary team with various expertise and skill sets. All participants stayed with the project for the full year.

Content was defined as what was supposed to happen in the process, specifically the activities we had to develop our educational game. The consensus was that due to the backward plan, “This team did complete the plan of work pretty much the way they had laid it out….”

Dose delivered included multiple core components including team meetings, wraparound services, wraparound service (NTAE staff) meetings, and a game jam. Our interviews provided insights on what activities team members most recalled or interacted with. For instance, a consensus was that “our team met once a month as a UConn team. And then we also had sometimes

subcommittee meetings." There were actually a variety of activities, "So remember we did both the game and the field book and webinars, so we actually had a lot of projects going on during the year, so it was time intensive." However, an overarching element of the project was,

> …he asked me to make what they were calling a backward plan, which…basically we took our end goal, and he asked me to plan what we needed to do each month to get to where we wanted to be in August 2021. So, I wrote that out and had and use that.

Dose received is the assessment of what program activities participants attended or were exposed to. For our context, it includes UConn team member meeting attendance, additional catalyst interactions, mediums used (online, email), and the time spent. A catalyst noted, "Just my experience … they were at every meeting, they didn't miss anything. They were actively engaged with us in email and other forms of communication." However, the feedback varied on the quality and effectiveness of each activity.

Quality refers to the degree to which personnel delivering the program did so in a way that achieves the intervention's stated objectives. Beyond assessing whether the intervention was delivered using the selected techniques, it can also encompass the level of training received by personnel. We surveyed team members on the perceived quality of both the project activities as well as team leader competencies and used three coding concepts in our assessment. The first was how productive or useful activities were in achieving the NTAE program objectives. Responses were mixed on the efficacy of the wraparound services, noting that "there were certainly some that were very useful, some that were probably a little less useful." Feedback stated that monthly check-in meetings were not as purposeful and useful as intended, and one process improvement is making meetings more effective or reducing their number. While meetings were "useful in terms of requiring accountability," and "they encouraged us," respondents felt that "we had too many meetings…some of the meetings basically were almost like formalities."

Conversely, this could be related to the structure of the meetings as often "everyone was …summarizing what had happened," and there was a concern that the push for consistency meant that there were no changes to the structure of the meetings until, "towards the end we tried to use that space differently." One other improvement suggested was not including all wraparound services at every meeting or having individual meetings with them. This suggests a downside to established check-in meetings if they are not designed in a way to encourage forward movement. The game jam received consistently positive reviews, noting that it was "the most useful part," and highlighting that, "you start to actually see a product and you start to see some forward movement and it's like ok, we are really doing something."

The second coding concept was the training given to personnel, in this case, the NTAE staff. In terms of training the catalysts received, UConn team respondents noted that "they're very experienced," and that wraparound service providers were "experts," and "all of our key informants have Extension experience." The third coding concept was whether there was enough or too much time. UConn team respondents suggested that "it would have been great to have more time with those wraparound services," because "we probably just didn't have enough time to fully engage to maximize them to that very useful level." Then again, one respondent suggested the program could have been more effective with a shorter time frame, stating that "in my view, the process could have been even these nine months we could've received all the help that we needed in a way."

Another way to rate the quality of delivery would have been to have an independent evaluator observe our meetings and leadership communications and rate them using an externally validated instrument. However, while potentially more informative, this would have been costly

and perhaps distracting. Each team must determine the best quality evaluation method using the context of their specific project.

Participant responsiveness concerns team members' perceptions of the relevance and utility of the NTAE process, as well as the degree to which they felt they actively participated. While ex-post interviews were used to assess this, it would have been better to also have mid-intervention points of data collection. We identified three main trends regarding participant responsiveness. The first was the importance of participative leadership. For instance, one NTAE staff noted that,

> [Team leader] was really good about coordinating with each one of the service providers like myself…. You know I said structure and clarity was was a strong point. That that allowed her and the team, … to take more advantage of different offerings from our team.

By actively engaging with both team members and external collaborators, team leadership ensured forward momentum on the project.

Career stage also impacts participative leadership, and the UConn team had members at varying levels. For example, one team member closer to retirement stated, "I'm happy to just lay low and be the support and do stuff in response to their needs…, lending support to junior and mid-career faculty."

The second trend concerned the role context and clarity played in participant responsiveness. For instance, one team member noted,

> So, it was really hard to be heavily involved in the game design piece because that is not an expertise that I have. So, there were sometimes, questions that I got posed that I was not in a position to respond to and that were better answered by individuals with that skill set and expertise. And same with some of the science or some of the other aspects that were more content focused….

However, conversely, a different team member suggested that, "I've participated more and in different ways than I would have if I'd had a clearly defined role." Personal interest also has an impact, as one participant noted that they "enjoyed the process so my interest kept me going and just the need to keep getting to these outcomes." A participant also suggested that, "what kept me involved was the pace of needing to finish the project," suggesting strict deliverable timelines can increase engagement. The third trend involved disparate participation. While a catalyst expressed that the team was putting in the required four to eight hours per week, some team members suggest this was not equally distributed across team members.

Looking at the monthly meetings specifically, it was suggested that while "people had various degrees of organization going into the meeting" as "there were times that some people were prepared and others were not," these forces balanced each other out. Thus, there may be a balancing act that leaders need to play, and one participant suggested that,

> I would say there should be a little bit more clear roles for people, so some people just say oh today I can't make it, oh tomorrow I can't make it, or I can be late. So, people, it's very easy for people to just not show up.

Based on the three trends we found, teams should consider the roles leadership, context, and clarity play in participation when developing program implementation.

While the project did not include activities specifically labeled as *facilitation strategies,* the semi-structured interviews uncovered activities that served this role of ensuring the team worked within the one-year NTAE timeline and met the program expectations. This began with initial project activities. First, every team developed a backward plan on working together as a team and with NTAE staff to meet their expectations. The logic of the backward plan is that "if

you don't begin with the end in mind if you don't put in some benchmarks that you're going to try to achieve along the way. You're probably not going to complete." Next, "each team has to put together what we call a unique value proposition," which requires them to articulate the goal and outcome for their project, and what differentiates it from other work in the space. Participants then complete their initial team health checks, a survey repeated multiple times.

The catalysts also had a conversation early on to "try to understand the strengths of each member of the team…and then how they can work together in productive ways." This leads to the second facilitation mechanism, which is the mentorship provided by the catalysts. Their "job is to ask powerful questions and to try to help leaders and teams think differently or more productively or more proactively, about their topic." Catalysts completed quarterly reports shared with all NTAE staff. It may not be possible to replicate this type of external mentorship within most Extension teams, but activities such as team health checks can be utilized. Additionally, "we did a kind of a vision casting session with the team where we asked, hey, where do you see this team in a year or two years…we would try to get some ideas and alignment on the team about where they're going." However, in terms of efficacy one UConn team member noted that they did not understand the surveys' purpose, so this must be clearly communicated.

Team meetings with the NTAE staff also played an important facilitation role. One catalyst suggested that,

> I think that is really useful because it brings all the wraparound services, key informant services, the team members, and other stakeholders into one place to provide updates and in those meetings, you usually gain a lot of information about the strategic direction of the team and where they are in terms of following their plan of work and accomplishing their goals.

In this way, even though the team meetings may have felt redundant to UConn team participants, they could lead to greater efficiency and synergies by preventing the silo effect that can happen if teams do not have information on other aspects of the project. One UConn team member suggested of the two monthly meetings that,

> Both were necessary. I think our team needed to meet internally to kind of figure out what our steps were to be ready for the meetings with the larger group to kind of talk about what we did. So, we needed both really both the internal meetings and the externals stay on task. Again, it was a 12-month project so it's quick. So, you have to meet frequently to get things done.

Meetings also allowed for reflection to "kind of step back and say ok, where are we at now? What do we really want to do? Do we need to pivot a little bit?" and, each meeting

> …would start with some recognition of the team members and how they had helped each other that past week or two weeks…so that there was recognition of the value that they brought to one another, and it reinforced their roles, their responsibilities, what they were good at.

As previously noted, a common refrain was that there were "too many meetings." This suggests that while team meetings can serve as a facilitation mechanism, this is predicated on their being organized so as to move the project further, rather than to primarily be used for summary of completed activities. The team leader's philosophy was inclusivity; there is no easy answer to the tradeoff between including all team members in all activities and limiting participation to honor other time commitments.

We defined *complexity* as the number of moving parts, uncertainty or confusion about activities, time available, and clarity versus ambiguity concerning activities and expectations. This

project was clearly one of high complexity. There needed to be organization and synergy of personnel from multiple institutions and fields, and there were a series of project activities and outcomes. For instance, our team at UConn was dependent on the team at NMSU for much of our project results, as well as the seven wraparound services provided by the NTAE staff. Not only were there many moving parts but there was also uncertainty on how to implement several of our activities. Thus, as with many Extension team projects, there was both high complexity and ambiguity.

The one-year time constraint was one area of complexity that never deviated. Concerning timing one UConn team member stated,

> So, we found out and then we just hit the ground running. If there had been a little bit more planning time internally in between that time, that probably would have been helpful so that we weren't scrambling to get the team together internally and get ready to work with them on the national level all at the same time.

While this was part of the process that the NTAE staff designed, and not one that our team could have changed, it does suggest that having a team and plan in place before beginning a project could decrease complexity.

One of the Foundation's learning objectives for the NTAE fellows and their teams was to innovate and increase their flexibility. Therefore, some ambiguity and the time constraint were deemed a necessary part of implementation for innovation to happen. A catalyst stated,

> And we're not expecting [the team leader] and the team to have everything worked out. What they did really well though, was have a plan to integrate at different times all of the wraparound services that we had available.

Thus, the NTAE staff recognized this complexity and helped teams mitigate it through the backward plan and other frameworks. Similar feedback on the experience with wraparound services was received, and one of the UConn team members stated, "More clarity and structure with the wraparound services might have allowed us to get more out of them." While some participants suggested clarity may have helped the UConn team feel more comfortable in the process, others noted it could have the unintended consequence of diminishing innovation.

We defined *context* as external implementation barriers and facilitators, which for our project included the pandemic and online environment, our Extension environment of being in different offices, time availability, funding, and the sequence of activities.

While the move to a virtual environment impeded certain activities such as in-person content development and user testing, it may also have potentially led to increased collaboration between personnel in distinct areas. For example, online meetings made it easier for all team members to participate regularly in meetings, even though they were located at different campuses and Extension field offices. Finding times that worked for everyone in a central location may have been prohibitively difficult. Additionally, while meetings with catalyst members would always have occurred online, as we were all working online for several months before the project began, our enhanced skills with online meetings and collaborations improved the project. The NMSU design team previously always worked in person, and we could not have worked with them without the pandemic, since our budget would not have allowed travel. Some participants agreed that the pandemic made completing the project virtually easier and more successful than if it had been attempted in-person. Conversely, several participants also mentioned that not being able to meet in person at all was challenging. Another key element of context is job demand. All personnel were involved in a variety of other activities and were facing challenges from COVID-related impacts. We did not realize the scope and scale of the NTAE project when the application was

submitted, and all team members agreed to participate. This project was more involved and time-intensive than anticipated, though team members felt it was important enough to warrant the extra commitments. However, we had the full support of UConn Extension's administrators. Another contextual factor is a significant funding constraint that led to a gap between our intended outcomes and what we could afford. We initially anticipated covering a variety of food labels in the game, in Spanish and English, with the assumption there would be sound and audio aspects. However, discussions with NMSU altered our plans, and instead, we selected the most important elements for an initial game because of budget and time constraints. As a result, the game prototype focuses on one food marketing label. Sound is limited to background music, and the game is only in English. Despite the alterations, the UConn team and NTAE staff still rated this project as highly successful, because of the team's cohesiveness, game produced, and additional funding received.

Discussion of Findings

There were multiple instances throughout the one-year project where a deviation or change could have led to very different project results, demonstrating why it is important to conduct a process evaluation. We developed four objectives for our analysis. First, we discovered that the eight process evaluation components all influence the impact and success of a project, and in the case of Extension teams, are interrelated. Some of these components were more influential as participants had more to say about complexity, quality, participant responsiveness, and context. Second, many activities can improve or hinder team dynamics. We found that meetings are important, though too many meetings can be burdensome. Communication is a crucial element impacting success, as is the development of a comprehensive project plan. Third, the activities of this particular NTAE project did accelerate the UConn team's development of a learning game and the team was successful in securing additional external funding to support programs. Fourth, we identified several areas where the NTAE and implementation structure improved team outcomes and leadership capacity.

Our first step in the process evaluation was creating questions for the semi-structured interviews and mapping them each to a specific evaluation category. We then developed our coding structure based on the definitions of each process evaluation category. Thus, in coding interviews, we looked at each interview holistically, rather than assuming responses to a specific question fell into the code that was originally assigned to that question. This led to revising our coding definitions as we began coding and realized there were overlaps between categories. For instance, originally "time" was related to both complexity and context. We realized we needed to further clarify that the one-year time constraint or the amount of time activities took should be coded to complexity, while the lack of time available to personnel belongs under context. We also realized that some questions, such as rating the quality of different activities, would have been much easier to code if we had given them actual scales to complete, rather than asking them verbally. This quantitative analysis could have then complemented our qualitative results. Once we began coding the interviews, we also realized that question responses did not always fit into our anticipated code, demonstrating the importance of coding holistically. In retrospect, certain process evaluation categories were better assessed through notes than interview responses. While participant opinions are crucial for understanding concepts such as quality, responsiveness, and context, we believe that content and dose delivered were best assessed using our backward plan. These notes contained extensive information on the original plan and the actual activities

implemented, and this would have reduced the burden of collecting data and coding for these categories. However, this is only successful in an environment with comprehensive notes.

Complexity is a moderator to fidelity and previous research shows that complex or vague interventions are at a greater risk for implementation issues (Hasson, 2010). It also leads to greater variability in delivery methods. The NTAE staff may have scheduled the monthly check-in meetings to help decrease any negative impacts. However, results were mixed on the efficacy of this choice. On the one hand, getting all involved parties together was useful in ensuring everyone was on the same page and that forward progress was occurring. However, the large number of participants also meant that much of the time was spent summarizing work completed, and there was often a subset of team members who did not feel a given discussion was relevant to them. Going forward, a better structure for these team meetings might be to send around updates the day before, and then have team members work together to address specific questions or next steps.

The complexity of the one-year time constraint could have been negative, but it appears to have positively impacted the project by forcing the team to be strategic and intentional with the development of their backward plan and their strategy for collaborating with the NTAE staff. Team members differed on the optimal amount of time needed for the project, potentially related to their perceived efficacy of activities, or the degree to which they were involved in all elements of the project.

Similarly, the context of COVID-19 served as a moderator throughout the process for the team and positively impacted both the process and outcome. The online contingency and one-year time constraint both could have been negative but were positive in this instance. Moreover, the limited funds, an important contextual element of this project, provided complexity. While it overall hurt the scale of our final game, it also forced the team to be creative and adapt in a way that allowed us to finish the project.

There was a high level of support for the team leader through meetings with catalysts, cohort sessions with other fellows, and one-on-one development sessions. The catalysts provided mentorship, ideas for the project and team, and overall strategy suggestions throughout the year. This support, along with that of the NTAE staff, helped ensure the team had high levels of participant responsiveness throughout the project. This correlates to the dose received, and teams must carefully consider the quantity and quality of the dose delivered, as we saw with meetings being rated as both useful and overdone. Quality and facilitation strategy must be considered alongside the dose delivered when designing a project. We used our backward plan and results from team health checks as formative evaluation tools, but there are also other ways to provide continuous measurement during project implementation. Moreover, the ongoing successes of the project can partially be attributed to the team's willingness to actively be open to new ideas and engagement.

Crucially, we did not assess whether there were competencies or personnel that were not included in the team and should have been, which could impact our conclusion of high reach. The team discussed adding more members and expertise three months into the project but then decided that an advisory board would provide the expertise needed without adding more complexity to the team. We also only considered reach from the standpoint of expertise, and future teams could consider demographics as well.

Experience is also a moderator and can impact program success. Facilitation and leadership experience both impacted this project, including its high rating. All interviewees mentioned both leadership from the team leader and leadership competency as highly effective. One catalyst noted that the fact that our team leader was a communicator significantly helped this team and aligns

with our earlier statement that communication is integral to successful program implementation. The leader's facilitation strategies should also be included as a moderating factor.

Limitations

Our analysis is not without limitations. First, many of the assessments occurred at the end of the project; this can be a limiting factor as it relies on participant memory. Our team started discussing a process evaluation soon after our project started, however without the framework, we did not begin implementing it until the project was completed. Ideally, we would have adapted the process evaluation framework at the start of the project and developed evaluation materials that could be used throughout the project year. Future process evaluations in social sciences should plan the process evaluation before starting the project and complete the process evaluation alongside the implementation. We compared the backward plan, or intended scope of work, to what actually happened as a formative evaluation for this project. Second, the interviews were conducted more than a year after project completion, and this meant that interviewees responded using memory, which can limit accuracy. Third, there was ambiguity in the interview coding related to whether the UConn team or NTAE staff was being referenced. Finally, the process evaluation team included two of the UConn team members, and this may have caused bias in some respondents, even though interviews were conducted by the evaluation specialist, who did not participate in the NTAE program with the team.

Conclusion and Recommendations

We performed a process evaluation of a one-year project to create a food labeling game. In conducting the process evaluation, by breaking up the team process into fidelity factors and moderators, we found that some of the components mattered more in our analysis, including complexity, quality, participant responsiveness, and context. Greater focus should be placed on these factors during program implementation and evaluation research should emphasize measuring and analyzing these factors.

There were areas where this program implementation could be improved, particularly around meeting purpose, meeting frequency, and content. In the end, a game prototype was successfully created that consumers play online, and we used to secure additional funding. Team dynamics and acceleration are strong, and the team is still working together. Thus, the intervention could overall be termed a success. There are two primary reasons for that. First, we found that the shorter time frame of this project created timelines that had to be met. Second, we also had the full support of the NTAE staff to help our project meet our goals.

One key element of success is that there were clear timeline expectations laid out by the NTAE program, but there was sufficient ambiguity to allow teams to adapt program activities to their specific conditions. Prior literature highlighted the importance of creating interventions with clear theoretical underpinnings and well-described non-negotiable elements to allow for local implementation variance that still ensures the spirit of the intervention is consistent (Carroll et al., 2007; Hasson, 2010; Moore et al., 2015). This was also true for the Extension teamwork. Recruitment was personalized, but without assessing whether any key competencies were missing from the team. Crucially, team dynamics are built around who is recruited and reached through the team, and the responsiveness of these participants. While recruitment and reach were given less emphasis in this paper, the authors acknowledge their importance in building strong team

dynamics. Team leadership also influences team dynamics, and that further enhanced this project, and needs to be considered in future process evaluations.

Small teams conducting a process evaluation may not have a lot of resources, and certainly, a full process evaluation can be daunting. Areas of focus may vary depending on the specific team, however, there are components they can focus on. First, starting the process evaluation before program implementation can save time. Second, we recommend measuring content and dose delivered using a backward plan and meeting notes, with notes maintained throughout the implementation. Semi-structured interviews or another qualitative approach should be used for quality, participant responsiveness, and context. Finally, complexity is hard to measure and an area where process evaluations are most beneficial so an essential component to measure. We used several measures, including time, number of people involved, and context; and also looked at collaboration and the number of partnerships.

We found that the modified process evaluation framework is effective when used in Extension, and likely in other social science realms, in addition to determining that our project was successful when measured against the third objective of accelerating our team. Process evaluations provide multiple outcomes, including developing leadership capacity among Extension professionals to support current and future initiatives, and including them in Extension programs could lead to greater overall impact. Clarity can have an unintended consequence of diminishing innovation (West et al., 2003), but our team did not explore this question in our process evaluation, and it is another area for future research. Future process evaluations could also incorporate formative check-ins throughout the process. Although using process evaluations in Extension is still relatively new, we expect more social science professionals to use them as program and funding landscapes become increasingly competitive.

References

Bull, N. H., Cote, L. S., Warner, P. D., & McKinnie, M. (2004). Is extension relevant for the 21st century?. *The Journal of Extension, 42*(6), Article 3. https://tigerprints.clemson.edu/joe/vol42/iss6/3

Carroll, C., Patterson, M., Wood, S., Booth, A., Rick, J., & Balain, S. (2007). A conceptual framework for implementation fidelity. *Implementation science*, *2*, 1-9.

Department of Health and Human Services. (2018, August). Developing process evaluation questions. *Centers for Disease Control and Prevention*. Available at https://www.cdc.gov/healthyyouth/evaluation/pdf/brief4.pdf

Dietrich, T., Rundle-Thiele, S., Kubacki, K., Durl, J., Gullo, M. J., Arli, D., & Connor, J. P. (2019). Virtual reality in social marketing: A process evaluation. *Marketing Intelligence & Planning*, *37*(7), 806-820.

Downey, L. H., Terp, K., & Byrd, S. H. (2021). A process evaluation of the HappyHealthy social marketing campaign. *Journal of Human Sciences and Extension, 9*(2), 9. https://doi.org/10.54718/PEJN1237

Duerden, M. D., & Witt, P. A. (2012). Assessing program implementation: What it is, why it's important, and how to do it. *The Journal of Extension, 50*(1), Article 5. https://doi.org/10.34068/joe.50.01.05

Duhigg, C. (2016, February 25). What Google learned from its quest to build the perfect team. *The New York Times Magazine*. https://centre.upeace.org/wp-content/uploads/2020/09/7.1-what-google-learnt.pdf.

Fink Chorzempa, B., Smith, M.D., & Sileo, J.M. (2019). Practice-based evidence: A model for helping educators make evidence-based decisions. *Teacher Education and Special Education*. 42(1): 82-92. https://10.1177/0888406418767254

Guevara-Villalobos, O. (2011, January). Cultures of independent game production: Examining the relationship between community and labour. In *DiGRA Conference* (p. 18).

Hamilton, N. (2019). Fostering and assessing law student teamwork and team leadership skills. *Hofstra Law Review*, *48*, 619.

Hasson, H. (2010). Systematic evaluation of implementation fidelity of complex interventions in health and social care. *Implementation science*, *5*(1), 1-9.

Holland, A., Troschinetz, A., & Missaghi, S. (2019). Get started or unstuck: Four elements of successful interdisciplinary collaboration in extension. *The Journal of Extension, 57*(5), Article 1. https://doi.org/10.34068/joe.57.05.01

Hulscher, M. E. J. L., Laurant, M. G. H., & Grol, R. P. T. M. (2003). Process evaluation on quality improvement interventions. *BMJ Quality & Safety*, *12*(1), 40-46.

Klug, M., & Bagrow, J. P. (2016). Understanding the group dynamics and success of teams. *Royal Society Open Science*, *3*(4), 160007.

Lamm, A. J., & Israel, G. D. (2013). A national examination of extension professionals' use of evaluation: Does intended use improve effort?. *Journal of Human Sciences and Extension, 1*(1), 4. https://doi.org/10.54718/UCLB3696

Lewis, C., Boyd, M., Walsh-Bailey, C., Lyon. A., Beidas, R., Mittman, B., et al. (2020). A systematic review of empirical studies examining mechanisms of implementation in health. *Implementation Science*. 15(1):21.

Linnan, L., & Steckler, A. (2002). Process evaluation for public health interventions and research. *Jossey-Bass/Wiley.*

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. *BMJ (Clinical Research Edition)*, *350,* h1258. https://doi.org/10.1136/bmj.h1258

Preston, J. A., Chastine, J., O'Donnell, C., Tseng, T., & MacIntyre, B. (2012). Game jams: Community, motivations, and learning among jammers. *International Journal of Game-Based Learning (IJGBL)*, *2*(3), 51-70.

Saunders, R. P., Evans, M. H., & Joshi, P. (2005). Developing a process-evaluation plan for assessing health promotion program implementation: a how-to guide. *Health promotion practice*, *6*(2), 134-147.

Tracy, S. J. (2019). *Qualitative research methods: Collecting evidence, crafting analysis, communicating impact*. John Wiley & Sons.

West, M. A., Borrill, C. S., Dawson, J. F., Brodbeck, F., Shapiro, D. A., & Haward, B. (2003). Leadership clarity and team innovation in health care. *The leadership quarterly*, *14*(4-5), 393-410.

Wiggins, G., & McTighe, J. (2005). Understanding by design (2nd ed.). Alexandria, VA: Association for Supervision and Curriculum Development ASCD. *Colombian Applied Linguistics Journal*. 19(140). https://10.14483/calj.v19n1.11490

Appendix 1. Elements of the selected process evaluation and the interview question items.

Component	**Definition**	**Sample of Interview Questions**
Recruitment	This refers to the procedures and program activities used to attract participants, including the resources that were employed and potential barriers to recruitment. Recruitment is an important moderating factor for reach, and for team assessment one would consider how team members were invited to participate, and whether certain collaborator subgroups were more or less likely to be successfully recruited. One should specify whether recruitment is through personal entreaties, general email requests, or some combination of both.	Do you remember how you were asked to join this team?
Reach	This is the degree to which the intended priority audience participates in the intervention and is often measured as a percentage of the total target population that participates. One might consider how many suggested collaborators chose to join or not join the team, especially for interdisciplinary open calls. A crucial decision is what value should serve as the denominator of your reach percentage, and one could include subgroup analyses to assess the diversity of recruitment and response.	What led to you joining the team?
Program content	These are the specific elements of the intervention, and include any drug, treatment, skills and/or knowledge a program is meant to deliver. In the case of an Extension team, "intervention activities" could correspond to planned steps undertaken in the process of creating programming materials or outreach projects.	Was a plan communicated or presented to you? Please describe the plan.
Exposure or dose delivered	This measures the frequency and duration of activities made available to participants. For a team evaluation, one must consider the degree to which project activities were implemented, the timeline was adhered to, and specified best practices were followed.	How would you rate your participation during monthly team meetings, relative to your expectations?

Component	Definition	Sample of Interview Questions
Dose received	This component refers to the extent to which participants actually receive and use the materials and resources implemented during the intervention. For an Extension team one could look at metrics such as meeting attendance or email open rates or take a qualitative approach to understand the degree to which team members felt they participated and contributed, were perceived by other team members to participate and contribute, or took advantage of available resources.	What kept you involved or engaging in the team?
Quality of delivery	This usually refers to the manner in which staff, doctors or teachers deliver a program, with assessment primarily concerning whether the intervention was delivered in a way appropriate to achieving intended outcomes, and whether that approach fits with the overarching theory. For a team setting, this could apply to leadership competency, meeting organization quality, and adequacy of resources provided to aid in program development. This is best assessed by an outside observer but can also be measured through participant interviews or surveys. In our evaluation, this construct was measured through the following observed variables: competency of the program catalysts (retired deans) and their training, leadership training given to the program facilitator or team leader, and resource adequacy.	To what extent did the NTAE program meet your expectations? Please explain.
Participant responsiveness	A key moderator of intervention fidelity, especially dose received, is whether participants respond to, or are engaged by, the intervention. It can be thought of as a measure of participant attitudes or enthusiasm, as well as whether they view the intervention as relevant to their experience and needs. This could be a measure of team engagement, either through quantitative survey instruments or qualitative focused discussions. Another important element could be the enthusiasm of the implementers, which could be thought of as the team leader (and also in our case the eight NTAE team members).	How useful did you perceive the wraparound services to be to the success of the project? How would you rate your participation during monthly wraparound meetings?

Component	**Definition**	**Sample of Interview Questions**
Facilitation strategy	This factor primarily consists of the support strategies employed to encourage implementation fidelity, including intervention monitoring, collection of formative feedback and use of training activities. While simple interventions may require few facilitation strategies, this becomes an increasingly important moderating factor as the degree of complexity increases. In our case, facilitation mechanisms included Team Health Checks (anonymous surveys throughout the project period on team member perceptions of team strengths and cohesiveness).	How do you rate team leadership on each of the following competencies?
Complexity	The complexity of an intervention can directly impact the degree to which it is implemented correctly. While complex (rather than simple) interventions are likely to have a higher level of delivery variance, this is also moderated by whether descriptions, instructions, and training were detailed or vague. Most Extension program development teams could be considered complex, though personnel have control over the degree to which expectations are clearly shared and the plan of action is detailed and described. We considered the time available for the project, the number and length of activities, and the expectations of team members.	How would you rate the level of project acceleration through NTAE?
Context	This consists of all external factors that serve as barriers or facilitators. Any understanding of behavioral change must account for the social system in which participants exist. Internally, this can involve their relationships and interactions with participating organizations and program implementers. Externally, one must account for environmental factors such as scarce funding, time constraints and most recently the COVID-19 pandemic.	How did COVID impact your experience, positively and negatively, in the project?

Appendix 2. Process Evaluation Codes.

Component	Codes
Content	● Advisory Support ○ Fellows meetings ● Team Meetings ○ UConn Extension Team (scheduled to be monthly) ○ NMSU State University ● Wraparound services ○ Digital audit ● Wraparound service meetings (planned to be monthly) ● Modified Game Jam (Intended to be one week)
Exposure/Dose Delivered	● UConn Extension Team meetings ○ Consistency in day/time ○ Number of meetings/activities invited to ○ One hour ● Wraparound meetings ○ Consistency in day/time ○ Number of meetings/activities invited to ○ One hour ● Wraparound service ○ Number of interactions with catalysts beyond meetings ○ Publishing meetings ○ Webinar/chat prep meetings ○ Leadership trainings ● Game Jam ○ Present Monday, Pitch Wed, Finalized Fri ■ They continued working ○ Meet after survey and presentation
Dose Received	● Meeting attendance ● "Off the books' catalyst interactions ● Where delivered (online/email) ● Time spent ● Interactions

Quality of Delivery	● Training given to catalysts/wraparound/NMSU ● Was there enough time/too much? ● How productive/useful activities were ● Sequence of activities ○ Does order of activities impact effectiveness of activities
Recruitment	● The plan to reach people ○ Ad hoc/networking
Reach	● The number of people actually reached ● Demographics of participants ○ Who did we not know about?
Participant Responsiveness	● Degree to which participated/was engaged in activities attended ● Enthusiasm, attitude
Complexity	● Number of moving parts ● Uncertainty/confusion about activities ● Clarity vs. ambiguity on activities vs. expectations ○ Logic model, backwards plan ○ One year time constraints ○ Level of project acceleration through NTAE
Facilitation Strategy	● Support strategies ● Team health checks ● Meeting check-ins ○ Intent of part of meeting ○ Did wraparound meetings actually serve this purpose? ● Leadership activities ● Meetings with fellows

Context	● Barriers ● Covid/online environment ● Extension environment—different offices ● Time ● Funding