



Reflections on experiential learning in evaluation capacity building with a community organization, Dancing With Parkinson's

April Nakaima^a, Sanjeev Sridharan^{a,b,*}

^a The Evaluation Centre for Complex Health Interventions, St. Michael's Hospital, Canada

^b University of Toronto, Canada



ARTICLE INFO

Keywords:

Evaluation capacity building

Realist evaluation

Brain health

ABSTRACT

This paper discusses what was learned about evaluation capacity building with community organizations who deliver services to individuals with neurological disorders. Evaluation specialists engaged by the Ontario Brain Institute Evaluation Support Program were paired with community organizations, such as Dancing With Parkinson's. Some of the learning included: relationship building is key for this model of capacity building; community organizations often have had negative experiences with evaluation and the idea that evaluations can be friendly tools in implementing meaningful programs is one key mechanism by which such an initiative can work; community organizations often need evaluation most to be able to demonstrate their value; a strength of this initiative was that the focus was not just on creating products but mostly on developing a learning process in which capacities would remain; evaluation tools and skills that organizations found useful were developing a theory of change and the concept of heterogeneous mechanisms (informed by a realist evaluation lens).

Through the Evaluation Support Program (ESP) of the Ontario Brain Institute (OBI), we had the privilege of being invited to work with multiple community organizations who deliver services to individuals with neurological disorders. As evaluation specialists our task was to help build evaluation capacity within those organizations. One of the features of this initiative was that the program of work was not overly prescribed; rather, we had the freedom and responsibility to co-construct along with the community organizations and to match the types of monitoring and evaluation approaches to what each community organization needed.

One key feature of this effort was its focus on theory-driven evaluation (Chen 1990; Rogers et al., 2000; Leeuw, 2003; Mayne, 2010; Coryn, Noakes, Westine, & Schroter, 2011; Leeuw & Donaldson, 2015). This implied getting each community organization to reflect on the multiple pathways by which it seeks to bring about impacts. Additionally our particular interest was thinking through the program theory of the community organization from a realist evaluation lens (Pawson & Tilley, 1997; Pawson, 2013). In that light we encouraged each of the organizations to think through the mechanisms that needed to be triggered along with the context and support conditions necessary for that particular intervention to work.

In this paper we discuss our learning about building evaluation capacity in the context of having worked closely with Dancing With Parkinson's, a charitable organization that has been offering weekly

dance classes to persons with Parkinson's disease for nearly 9 years at numerous locations in Toronto, Canada. Our response for this forum is organized into these central questions: 1) What have we learned about building evaluation capacity as a result of this initiative? 2) By what mechanisms can such evaluation capacities be built? 3) Based on our experiences, what have we learned about potential improvements to the ESP model of pairing evaluation specialists with community organizations to work closely together?

1. What have we learned about building evaluation capacity? What does it take to build capacity of an organization?

Unlike some of the other capacity building experiments that we have been a part of, this initiative was not defined purely as a training exercise. Instead, over a defined period of approximately 10 months, the evaluation team worked closely with the community organizations to better understand their needs. Building relationships was key in our implementation of the model discussed by Arasanj and Nylen (this forum). Evaluation capacity building is often discussed in the literature along the lines of training and actions employed to develop knowledge, skills and attitudes within organizational staff to be able to conduct program evaluations, use findings and evidence to improve programs and inform decision-making, and increase the likelihood of achieving intended results (Stockdill, Baizerman, & Compton, 2002;

* Corresponding author at: The Evaluation Centre for Complex Health Interventions, St. Michael's Hospital, Canada.
E-mail address: sanjeev@msn.com (S. Sridharan).

Carman & Fredericks, 2010; Labin et al., 2012; Preskill & Boyle, 2008; Fraser, Tremblay, Bourgeois, & Lahey, 2013; Cousins, Goh, Elliott, & Bourgeois, 2014). Often the choice of words in the literature hints at assumptions that program staff does not use evidence, that programs need improvement, and that desired outcomes are not being achieved. What we found with Dancing With Parkinson's is that the developers of the program have systematically, through attentive observation, dialogue with program recipients, staying abreast of research findings and consulting with neurologists, developed, "tested," and refined their program and service delivery. We found that the DWP program had been exceeding expected outcomes from the perspective of program recipients. One indicator was that the vast majority of clients have continued attending once they tried it, with some having attended week after week since the program started nine years ago. Retaining the vast majority of participants for even one year is a significant achievement – research has shown that 50% of adults who start a formal exercise program drop out by 6 months (Desharnais, Bouillon, & Godin, 1986; Dishman, 1991; Ettinger et al., 1997). Despite such great indicators, DWP struggles to stay afloat. The purpose for evaluation in the case of DWP became one of evidence/knowledge generation: systematically collecting evidence to prove that dancing is having a profound positive effect on program recipients and demonstrating the value of the organization. Community organizations are under pressure to demonstrate their value to funders in order to sustain their programs. Given that community organizations often are struggling to survive despite excellent evaluation findings, sustainability of evaluation capacities within an organization becomes a mute point if the organization itself is not sustainable. The "Catch-22" is that the need for evaluation becomes more imperative when resources are more constrained, but the limited resources restrict the sustaining of evaluation in favor of sustaining the program.

As mentioned above, relationship building was important for this collaborative model of evaluation capacity building. Our original plan was to connect with colleagues from Dancing with Parkinson's once a month. However we soon found these conversations around pathways of change, measurement, and pathways of influence (Mark & Henry, 2004) both challenging and exciting that we soon began to meet weekly over cups of coffee.

One advantage of having a structured 10-month plan was that the structure helped build momentum towards a clear end point. Specific products that we worked on included developing theories of change, developing a design to study the impact of DWP, and developing data collection protocols. Each of the steps were done in a highly collaborative manner. One of the real strengths of this initiative was that the focus was not just on creating products but also on having a learning process in which capacities would remain with the organization while working on an evaluation. Ten months is not much time to collaboratively design and complete an evaluation and accomplish capacity building with individuals from community organizations who already are quite busy and often over-extended; however, the limited time frame pushed us to look for opportunities and to be creative in designing a feasible evaluation while realistically putting into place simple monitoring ideas, plans and protocols for future evaluation strategies that would remain with the organization.

It is worth reiterating that DWP is a small organization that struggles with issues of sustainability. (For example, unlike therapies such as physiotherapy that are often covered by extended health care plans, dancing is not covered so clients pay for classes, and DWP fundraises so that clients pay less for classes). This project served as a catalyst for evaluative thinking with the key leads of DWP. Many of our dialogues with the key leads of DWP focused on issues of sustainability and how evaluations can help create an evidence base that can help with issues of sustainability. Just as pharmaceutical companies systematically collect evidence that a drug makes a difference in a patient or population group's health trajectory, evidence that community interventions such as DWP can have a substantial impact on clients' health

trajectories needs to be documented. DWP would like to show that at the very least dancing improves clients' quality of life, but also improves functioning, as clients report, and also may possibly slow down the degeneration associated with Parkinson's Disease.

We were strongly encouraged to have a steering committee that could guide the evaluation along. Given the small size of the DWP team, we were a little hesitant to have a steering committee guide the process as we depended on informal meetings, often arranged opportunistically to build momentum for this evaluation capacity building. However, we have seen steering committees serve a very useful function in other projects. We think that a steering committee does help raise the salience of an evaluation with a larger organization; however, we are not persuaded that a steering committee does not interfere with more informal or spontaneous processes of learning that might be more expedient for an organization to learn evaluation skills and knowledge. Our point here is that enthusiasm and excitement for the evaluation needs to be developed between the organizational staff working on the evaluation and the evaluation team.

Another strong feature of this capacity building initiative was its attention to geographical proximity between the evaluation team and the community organization. The proximity to the community organization helped with arranging informal meetings and as stated above taking a more opportunistic pathway to knowledge building.

The planning for this initiative was a strength, and in particular planning for meetings attended by all organizations, with five different organizations chosen in the first year and paired with two teams of evaluation specialists. As the evaluations progressed, the teams were asked to present their thinking in full-day meetings. These meetings were arranged three times over the course of the ten or so months. We think there was value in such meetings for learning about what other teams were doing, how they were approaching various challenges, and also in building a potential community of organizations interested in issues of evaluation.

2. Understanding mechanisms

It is useful to reflect on the multiple mechanisms by which an initiative such as the Ontario Brain Institute's Evaluation Support Program can help build capacities.

2.1. Attitude towards evaluation

As noted earlier, we think there is value in building the confidence and appreciation of community organizations in conducting evaluations. In our experience most community organizations have had very negative experiences with evaluation. Even the notion of evaluative thinking is alien and difficult for most organizations. First and foremost we think that an initiative such as this one changes the mental space around evaluations. The idea that evaluations can be friendly tools in implementing meaningful programs is in itself an important insight and one key mechanism by which such an initiative can work.

2.2. Relationships

Second, relationship-building between the community organization and the evaluation team, in our judgement, is quite pivotal in the success of such an initiative. As noted earlier, in our experience with Dancing with Parkinson's, some of the most important learning moments happened outside of the formal meetings. While it was not our original plan to meet more than monthly, we found it both exciting and useful to meet on a close to weekly basis. We think that such relationship-building helped raise the enthusiasm for evaluation and helped change an implicit view of how useful evaluations can be for small community organizations. This initiative of Dancing with Parkinson's formally ended about one year ago, but we continue to meet both socially and professionally to further our dialogues.

2.3. Useful tools of evaluative thinking

Third, a key mechanism was around key tools and skills that evaluation has to offer. We found our discussions around theories of change powerful in framing the evaluation. As we worked on the theory of change there were multiple “ah-hah!” moments as key planners of DWP recognized that the pathways by which their intervention impacted individuals were more diverse than they themselves had considered. Another concept that we found helpful in our dialogue in thinking theoretically was around heterogeneity, specifically around heterogeneous mechanisms. We stressed in our dialogues that different individuals with different contexts might need different program mechanisms for outcomes to be impacted. This thinking was informed from a realist lens, and it was powerful in both formalizing the program planning aspects of DWP and also helped inform the evaluation design. Discussions around monitoring and evaluation design were also helpful in building the skills of our partners at DWP.

However, one area that we were less successful with was analysis. By the time the data was collected, the ten months had elapsed and we did not have enough time to follow individuals for the required period of time. However even this challenge worked as a learning device for our partners to appreciate that understanding the value added of specific interventions in terms of its impacts on individuals takes time.

2.4. Learning from other groups

A fourth mechanism by which this initiative potentially worked was creating spaces in which all the teams would meet periodically to discuss progress in conceptualizing and operationalizing the evaluation. As noted earlier, three such meetings were planned. These meetings helped create a community of practice around evaluation. We felt a good insight from the planners of this initiative was not to treat capacity building purely as a series of products but as a dynamic learning process. From our own experience, teams enjoyed learning from each other about their respective theories of change and how they planned to institute measures of progress and performance. They also were encouraged, in the sense of their confidence being boosted, and delighted in other groups asking how they managed to address a problem, thus sharing strategies. For example, a discussion arose at one workshop around how best to elicit feedback from clients who experience cognitive challenges like short-term memory loss – the Alzheimer’s Society group found that it’s preferable to get feedback right after a client engagement rather than doing a seasonal evaluation (like surveys distributed to clients once a year).

3. Lessons learned: potential improvements to the model

In an evaluative spirit, we also think that our experience working with Dancing With Parkinson’s and other capacity building projects as part of the OBI initiative has highlighted a few critical questions that we think can help further build the evaluation capacity building field.

First, is one phase enough for building organizational capacity building? We start with recognition in a fiscally constrained environment that it is remarkable that an organization such as OBI is raising the dialogue around building evaluation capacities and evaluation culture in community organizations. As discussed earlier this initiative is intended to last for close to 10 months. We recognize that a number of capacity building projects are even shorter in duration, often little more than a couple of weeks of training. We still think that in a spirit of building a field, 10 months may not suffice to build the organizational capacities for organizations such as Dancing with Parkinson’s. What Phase 1 does is build enthusiasm and support for evaluation and evaluative thinking among key stakeholders of an organization. It also introduces them to the importance of thinking evaluatively in helping make claims for the value added of programs and organizations. However, we don’t think that 10 months suffices to build capacities in a

sustained way within organizations. We also think that given chronic conditions like Parkinson’s disease, establishing impacts is not a short-term problem. As an example, in our work while we were able to collect a range of data we were unable to implement the entire suite of instruments to establish causation. Establishing causation is especially challenging given the heterogeneous needs of persons with Parkinson’s. As example, the individuals we interviewed were at varying stages of Parkinson’s, and impacts of an intervention might mean different things for different individuals. For all the reasons outlined above, we think there is value in a second phase for such initiatives.

Second, there’s also value in raising more focused questions around what it takes to sustain capacity building efforts. Eventually an organization like OBI can provide support for capacity building initiatives in a time-limited way. One can argue that even 10–11 months is an impressive amount of time for support. That noted, we think there is value in such a forum to raise questions around what it takes to sustain such initiatives. A comment made by Jean King in multiple dialogues (personal communication) was that a number of evaluation capacity building efforts do not pay attention to issues of sustainability. Going forward a critical question has to be: How does one plan for sustainable capacity building? We think a focus on the value added of a specific capacity building model is less important than an explicit focus on sustainability.

Third, in addition to DWP, as a team we have felt lucky to work with multiple organizations. These have included: Epilepsy Toronto, Suvien (app for Alzheimer’s), N’mninoeyaa (First Nations disabilities project), and Three To Be. What has been impressive in our experiences working with these multiple organizations is that we have felt that there may be a relationship between organizational size and the capacity building model. As example, one of our learnings was that a larger organization like Epilepsy Toronto might need a different type of capacity building support than a smaller organization like DWP. We also feel, given the pivotal role that relationship building plays in capacity building efforts, organizational size might impact such relationship building efforts.

Fourth, we think as a field it is also important to develop clearer metrics for what constitutes progress in capacity building efforts. As an example, we think having measures, both qualitative and quantitative, of relationships is important, but it is equally important to think of other measures of capacity at the organizational level including rigor, organizational skills, organizational support and organizational level knowledge. Further, exploring measures (either qualitative or quantitative) that measure the organization’s long-term commitment to evaluation capacity as well as measures for evaluation culture might be needed.

Fifth, as a field we think we need better stories of the dynamics of organizational capacity building from specific case studies of organizations. Specific questions for framing such case studies might include: How long did it take for specific organizational capacities to be developed? What mechanisms helped build and sustain such evaluation capacity building? What happened when key champions of evaluation left such organizations? What were platforms that were developed to help pass the baton between different staff during times of transition?

Sixth, working on this initiative has also helped us reflect on what types of skills are needed to build capacities in complex community organizations that are providing a range of different services that respond to individuals’ needs and trajectories of the disease. It is not simply a matter of one size does not fit all, but community organizations have to pay attention to the trajectory of a specific individual’s disease – one size does not fit a single individual’s needs considered dynamically over time. This point has implications for methodology, and we reflect on the methodological import of this point in the paper by Sridharan and Nakaima in [this forum](#).

4. Conclusions

We return to the domain issues of neurological disorders. Part of the

ambition was to provide greater incentives for community organizations that play a pivotal role in creating a community of care for individuals and families who are impacted by neurological disorders. We applaud OBI for taking the lead in framing a broader dialogue around creating greater incentives for organizations like DWP to build an evaluation culture. It is worth noting, however, that the asymmetries of the capacities that pharmaceutical companies and more importantly the incentives that such companies have to build an evaluation culture around pharmaceutical products are considerably greater than what exists in community organizations. We think that OBI has done a tremendous service in raising the dialogue around raising incentives for community groups. We think other organizations also need to step up in such a dialogue to help continue to provide incentives for community organizations. As noted, the asymmetries and incentives are great and the journey towards building capacities for the community sector is a long and steep one: more partners need to join the table.

References

- Carman, J. G., & Fredericks, K. A. (2010). Evaluation capacity and nonprofit organizations: Is the glass half-empty or half-full? *American Journal of Evaluation*, 31(1), 84–104.
- Chen, H. (1990). *Theory-driven evaluations*. Newbury Park: Sage Publications.
- Coryn, C. L., Noakes, L. A., Westine, C. D., & Schroter, D. C. (2011). A systematic review of theory-driven evaluation practice from 1990 to 2009? *American Journal of Evaluation*, 32(2), 199–226.
- Cousins, J. B., Goh, S. C., Elliott, C. J., & Bourgeois, I. (2014). Framing the capacity to do and use evaluation. *New Directions for Evaluation*, 141, 7–23.
- Desharnais, R., Bouillon, J., & Godin, G. (1986). Self-efficacy and outcome expectations as determinants of exercise adherence. *Psychological Reports*, 59, 1155–1159.
- Dishman, R. K. (1991). Increasing and maintaining exercise and physical activity. *Behavior Therapy*, 22, 345–373.
- Ettinger, W. H., Jr., Burns, R., Messier, S. P., Applegate, W., Rejeski, W. J., Morgan, T., et al. (1997). A randomized trial comparing aerobic exercise and resistance exercise with a health education program in older adults with knee osteoarthritis. The Fitness Arthritis and Seniors Trial (FAST). *JAMA*, 277, 25–31.
- Fraser, M. A., Tremblay, G. H., Bourgeois, I., & Lahey, R. (2013). Neutral assessment of the national research council Canada evaluation function. *Canadian Journal of Program Evaluation*, 28(2), 85–96.
- Labin, S. N., Duffy, J. L., Meyers, D. C., Wandersman, A., & Lesesne, C. A. (2012). A research synthesis of the evaluation capacity building literature. *American Journal of Evaluation*, 33(3), 307–338.
- Leeuw, F. L., & Donaldson, S. I. (2015). Theory in evaluation: Reducing confusion and encouraging debate. *Evaluation*, 21(4), 467–480.
- Leeuw, F. L. (2003). Reconstructing program theories: Methods available and problems to be solved. *American Journal of Evaluation*, 24, 5–20.
- Mark, M. M., & Henry, G. T. (2004). The mechanisms and outcomes of evaluation influence. *Evaluation*, 10(1), 35–57.
- Mayne, J. (2010). Contribution analysis: Addressing cause and effect. In R. Schwartz, K. Forss, & M. Marra (Eds.), *Evaluating the complex*. New Brunswick, NJ: Transaction Publishers.
- Pawson, R., & Tilley, N. (1997). *Realistic evaluation*. London: Sage Publications.
- Pawson, R. (2013). *The science of evaluation: A realist manifesto*. London: Sage Publications.
- Preskill, H., & Boyle, S. (2008). A multidisciplinary model of evaluation capacity building. *American Journal of Evaluation*, 29(4), 443–459.
- Rogers, P. J., Petrosino, A., Huebner, T. A., & Hacsí, T. A. (2000). Program theory evaluation: Practice, promise, and problems. In P. J. Rogers, T. A. Hacsí, A. Petrosino, & T. A. Huebner (Vol. Eds.), *Program theory in evaluation: Challenges and opportunities, New Directions for Evaluation*, vol 87, (pp. 5–14). San Francisco: Jossey-Bass.
- Sridharan, S., Zhao, K., & Nakaima, A. (2017). Editors' notes. In S. Sridharan, K. Zhao, & A. Nakaima (Eds.), *Building capacities to evaluate health inequities: Some lessons learned from evaluation experiments in China*. Jossey-Bass: New Directions for Evaluation, Number 154.
- Stockdill, S. H., Baizerman, M., & Compton, D. W. (2002). Toward a definition of the ECB process: A conversation with the ECB literature. *New Directions for Evaluation*, 93, 7–25.