Practice Article

EJA EVALUATION Journal of Australasia

Evaluation Journal of Australasia
2020, Vol. 20(1) 6–22

DOI: 10.1177/1035719X20909910
journals.sagepub.com/home/evj
SAGE

Lessons learned in evaluating the infrastructure of a Centre for Translational Research

Ralph Renger
JUST Evaluation Services, USA

Marc D Basson, Gary Hart, Rick Van Eck, Eric Souvannasacd
University of North Dakota, USA

Jessica Renger
Claremont Graduate University, USA

Jirina Foltysova
Walden University, USA

Abstract

This article shares lessons learned while evaluating the implementation of a Clinical and Translational Research Centre (CTR). To meet its overarching goals, the CTR consists of numerous research support units (e.g., biostatistics, community engagement, professional development) that are intended to work together collaboratively. It is then argued that an evaluation approach grounded in system thinking was the best fit to evaluate this key CTR design feature. The rationale for selecting systems evaluation theory (SET) as the evaluation framework best suited to evaluate the CTR infrastructure is then presented. The application of SET and the lessons learned are then shared. This article concludes that there are many similarly

Corresponding author:
Ralph Renger, JUST Evaluation Services, LLC, 14777 E. Circle M Ranch Pl, Vail, AZ, 85641, USA.
Email: ralph@justevaluation.com

structured programmes worldwide to which the lessons learned can be applied and upfront investments in using a system approach are rewarded by providing meaningful and useful evaluation recommendations for system change.

Keywords
evaluation, logic models, systems evaluation theory, systems thinking

This article shares lessons learned while evaluating the implementation of a Clinical and Translational Research Centre (CTR) funded by the Department of Health and Human Services (DHHS) National Institutes of Health (NIH). According to the NIH (2017), CTR funding is primarily intended to '. . . enhance the competitiveness of the CTR trained investigators [in the CTR] to obtain additional funding for clinical and translational research' (p. 2). To meet this goal, the NIH mandates that each CTR develops a supporting infrastructure comprised of several key component activities (KCAs), or cores, including (1) Professional Development; (2) Biostatistics, Epidemiology and Research Design; (3) Community Engagement and Outreach; (4) Pilots Programme (i.e., studies with the potential to grow into independent research programmes); (5) Administration (overall CTR leadership and organization as well as such activities as purchase and sharing of lab equipment across CTR collaborating institutions) and (6) Tracking and Evaluation (TAE).

The Dakota Cancer Collaborative on Translational Research (DaCCoTA) concentrates on advancing cancer research in rural populations as well as American Indians and other minorities. The DaCCoTA CTR organizational structure is shown in Figure 1. The CTR organizational structure mirrors that of many government-funded initiatives worldwide. The Australian health infrastructure grants (https://www.nhmrc.gov.au/research-policy/research-translation/recognised-health-research-andtranslation-centres), the Canadian National Research Council Collaboration and Research Centers (https://nrc.canada.ca/en/research-development/research-collaboration/), the Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) Germany (https://www.giz.de/en/worldwide/germany.html) and initiatives in the United Kingdom, including (https://mrc.ukri.org/research/initiatives/) and (https://www.england.nhs.uk/ourwork/part-rel/ahsn/) are examples of where governments provide infrastructure funding to support the coordination of research support units. Thus, it is our hope the lessons learned and shared here are of utility to evaluators worldwide.

In addition to providing clinical and translational researchers with the support necessary to achieve independent funding, all CTRs have a second overarching goal to move projects along the translational research continuum (see Figure 2; Institute for Clinical and Translational Research [ICTR], 2019; Rubio et al., 2010).

The exact organizational structure of other CTRs may vary slightly, and some CTRs add 'optional KCAs'. However, CTRs generally have similar educational and service objectives that include professional development and mentorship, familiarity with principles of clinical and translational research, an understanding of ethnic cancer disparities related to CTR geography and behaviour, creation of a system by which

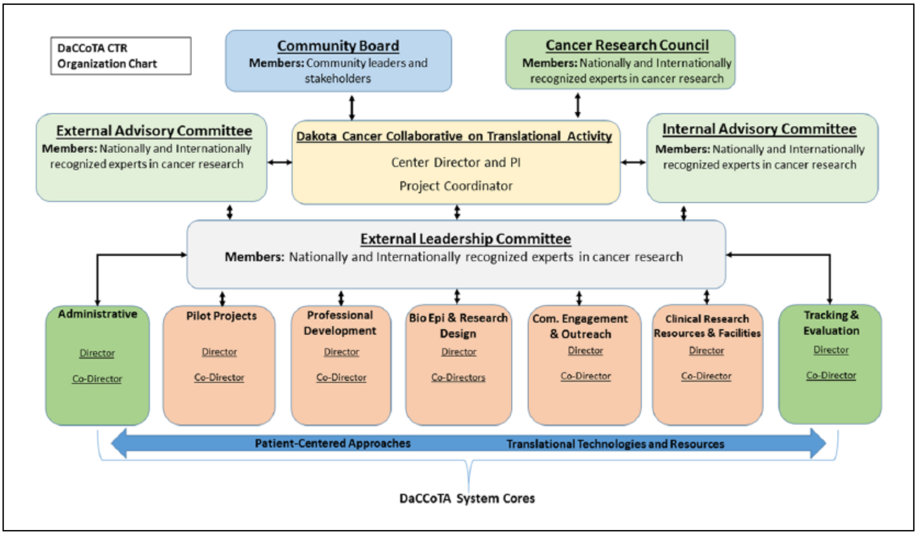

Figure 1. DaCCoTA organizational chart.

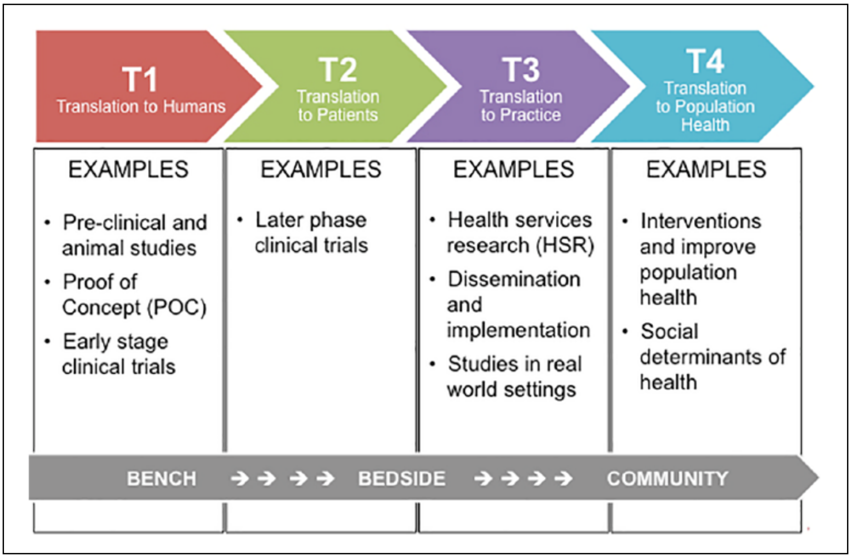

Figure 2. The Cancer Research Translational Continuum (Rubio et al., 2010).

resources can be made available for cancer study and treatment within the CTR region and facilities and systems of training CTR researchers in appropriate grantsmanship (e.g., sharing of services and equipment; DaCCoTA, 2019).

The purpose of this article is to share our Training and Evaluation (TAE) core[1] challenges, solutions and lessons learned in evaluating our DaCCoTA CTR. Our TAE core is mandated to provide evaluation support at two levels: at the project/investigator level and at a centre (i.e., infrastructure) level. The focus of this manuscript is on the evaluation of the DaCCoTA infrastructure in producing independent investigators and moving cancer projects along the translational continuum. We focus specifically on lessons learned in evaluating the efficiency and effectiveness of the CTR infrastructure in meeting these two goals.

Evaluating CTR infrastructure efficiency

To find the best suited approach for evaluating efficiency, we began by considering the key design features of the CTR core infrastructure. We needed to find the right key to open the lock (Williams, 2010). We reasoned an evaluation approach grounded in systems thinking was necessary because the CTR consists of several cores that are intended to work collaboratively towards meeting the common goals. That is, the CTR cores are supposed to operate as a system. The American Evaluation Association (AEA, 2018) defines a system as '. . . a set of interrelated elements that interact to achieve an inherent or ascribed purpose (Ackoff, 1971; Meadows & Wright, 2008)' (p. 6).

Our search for a systems-based evaluation approach led us to adopt systems evaluation theory (SET; Renger, 2015). SET emphasizes using systems thinking and systems concepts to evaluate interconnectedness; the essential CTR design feature. Furthermore, at the time of this writing, SET was the only published *evaluation* theory specifically designed for guiding evaluators in evaluating systems.

SET consists of three steps: (1) defining the system, (2) evaluating system efficiency and (3) evaluating system effectiveness. Step 1, defining the system, also consists of three steps: (1) establish the system boundaries, (2) identify the system elements and (3) detail the within and between system element relationships. The three steps to defining a system are grounded in the AEA systems thinking principles (AEA, 2018) and are purposively sequenced to continually drill down to define system details: akin to building a puzzle by beginning with the border (boundaries) and then filling in the remaining puzzle pieces (elements and relationships; Renger et al., in press).

To define the system boundaries, SET recommends first defining the system goal and then determining who shares this goal (Renger, Foltysova, Renger, & Booze, 2017). Recognizing the importance of goal clarity (Stazyk & Davis, 2019), we revisited the NIH (2017) guidance and defined the goal specifically as,

> to provide added value to the biomedical research efforts in the participating institutions through support of activities that cannot easily be provided through standard research grant awards. The proposed activities will provide the infrastructure and resources that will enhance the competitiveness of the investigators to obtain additional funding for clinical and translational research. Applicants must establish a statewide network of collaborating and partnering institutions/organizations. Other institutions/organizations outside the state may be included if forming a network of wider reach. (p. 2)

Our CTR leadership (i.e., the principal investigator and co-investigators) began by first examining the extent to which NIH-specified project support (i.e., core expertise), resided within our lead institution. After completing an internal search, the scan for core expertise expanded to include surrounding institutions in the Dakota geographic proximity. It should be noted that our CTR is also based on notions of collaborative team science, and the CTR leadership therefore made a philosophical decision to distribute the cores across the network of institutions and to further distribute leadership for each core among the network, even if sufficient expertise was present within a single institution. While this posed communication and coordination challenges, leadership viewed this as symbolic of the desired collaborative network structure.

The NIH guidance allows CTRs to expand geographical boundaries as long as they do not infringe and compete with existing CTRs. In this way, establishing the CTR system boundaries was both a conceptual and geographical endeavour. As a result of this process, CTR leadership established subcontracts with 13 institutions across two states. As per NIH guidance, the expertise was organized into KCAs, or cores, with an assigned director and co-director for each core who were responsible for overseeing staffing and operations. In systems thinking terminology, each CTR core is a system element.

The final step in defining the CTR system was to engage the core leaders in a process flow mapping (PFM) exercise (Renger et al., 2016). The purpose of the PFM was to detail each core's standard operating procedure (SOP). The SOP details the relationships within and between system elements.

The PFM exercise was facilitated by one of our evaluation core members and took between 60 and 90 min to complete. The initial PFMs were conducted independently with each core. This was intentional to allow the leaders of each core to thoughtfully consider their own processes. SOPs were then validated using a member check (Birt et al., 2016; Hoffart, 1991; Lub, 2015; Renger & Bourdeau, 2004). This was done by providing core directors a visual and written summary of the meeting and asking them to reflect on their accuracy. For illustrative purposes, a portion of a validated SOP from one of the CTR cores is shown in Figure 3.

We then brought core directors together in a tabletop exercise type format (United States Department of Homeland Security, 2020). To do this, we presented core directors with a hypothetical scenario of onboarding a new CTR project. As the scenario evolved, core directors had an opportunity to reflect on what changes they needed to make their SOPs, but just as importantly, where their processes intersected with other cores and the changes they needed to make to accommodate them.

Once the CTR system boundaries, elements (i.e., cores) and relationships (i.e., SOPs) were defined, we were prepared to evaluate CTR system efficiency. We began by evaluating each core's SOP. The core SOPs provided the standard of acceptability for the evaluation (Green et al., 1980; Renger et al., 2016). In other words, the SOPs describe what is supposed to be done, by whom and by when. We evaluated the extent to which the SOPs were implemented with fidelity using three primary strategies, described in the following sections.

Comparing actual SOP processes to validated SOPs

First, the steps core staff actually took in supporting a project were compared to the validated SOPs. We did this by conducting structured interviews with both investigators and core staff members at the end of the pre-award, implementation and close-out phases (London & Beatty, 1993). We opted for this strategy for several reasons. First, we wanted to gain needed perspectives from both those providing and receiving CTR support (AEA, 2018). Second, we were concerned waiting until a project was completed: (1) would delay feedback that could improve CTR efficiency for new projects and/or (2) might lead to poorer data quality because of subject memory decay.

Deviations from the validated SOP signalled one of the two corrective actions. If the SOP deviation was one of omission, then we recommended core leadership to reinforce the importance of adhering to the SOPs and leveraging the resources needed for additional training.

If the SOP deviation was one of the commissions, then we determined whether the deviation was an inefficiency or an efficiency. Renger et al. (2018) note that in some instances, an observed SOP deviation may actually represent an organically developed *improvement* in system efficiency. That is, the SOP deviation may represent a solution to an existing system inefficiency derived out of necessity by core staff. If the SOP deviation is in fact deemed to be an organically derived fix, then it signals that there is a disconnect between the core leadership (who developed the validated SOPs) and the core staff that needs to be addressed. In short, it may be the difference between theory and practice.

Examining SOPs for reworks

Reworks occur when SOP steps are repeated. They are a signal of potential system waste (Nave, 2002; Renger et al., 2018). When a rework was discovered, we met with core staff to understand the reason for the rework and suggest solutions to avoid system waste (Arnheiter & Maleyeff, 2005; Bentley et al., 2008). For example, we learned that the process for obtaining human subjects' approval from internal review boards (IRBs) was being reworked numerous times. At best, projects were being submitted for IRB approval at two or more institutions: the IRB of a CTR affiliate institution and the lead institution IRB. In some cases, such as projects involving Veterans Affairs (VA) or collaboration between institutions (e.g., hospital and university or even multi-hospital or multi-university studies), three or more IRB approvals were being required. Passing the same project through multiple IRB boards for approval drained finite system resources and delayed project start dates. For these reasons, our TAE core recommended seeking a blanket IRB approval, either through a private third party or by establishing a series of memorandum of understandings (MOUs) between all the partner institutions.

Evaluating SOPs using system concepts

SOP inefficiencies were also evaluated by applying several system concepts, including feedback loops, cascading failures and reflex arcs. Each system concept offers a different way of analysing connectedness between system elements.

Feedback loops. Systems require ongoing feedback to make necessary corrections to changing internal and external conditions (Banathy, 1992; Renger, 2016). We examined SOPs to identify internal (i.e., within cores) and external (i.e., between cores) feedback loops.

Each feedback loop was first evaluated to be sure it was closed. Critical information needed for CTR improvement must be passed on to the intra- and inter-core staff who need it. For example, in Figure 3, we noted that a feedback loop was not being closed after the recommendation for project funding. Those reviewing funding applications were under the assumption that they were to send their feedback directly to the applicant. However, the agreed upon process was to first integrate reviewers' comments into a single document, thereby protecting reviewer anonymity. The integrated reviewer feedback would then be shared with the applicants. Corrective actions were immediately taken to educate those responsible for providing feedback as to the agreed upon SOP and the problem was remedied.

Once a feedback loop was confirmed as being closed, feedback quality was evaluated. We used five criteria to evaluate feedback quality: *credibility, relevance, specificity, timeliness* and *frequency* (Choi & Myung, 2017; Goodwin & Miller, 2012; United States Office of Personnel Management, 2019).

Credibility. For CTR staff to act, feedback must be viewed as credible. Credibility is directly related to the feedback source. That is, a system actor is more likely to act upon feedback if he or she recognizes the source and/or has a relationship with him or her (Ilgen et al., 1979). Early in the CTR evolution, core directors were unfamiliar with their counterparts and in many cases did not know each other's names. North and South Dakota are two sparsely populated adjacent states that when combined cover approximately 185,000 square miles. Many of the Dakota medical facilities and personnel are great distances from one another. Failing to act on a request for assistance because the source was not deemed credible could lead to significant delays and potentially create a culture of mediocrity. To address this potential problem, monthly CTR-wide meetings and the participation of core personnel in CTR working groups helped core directors become better acquainted with their counterparts within and between institutions. The CTR director also required all CTR leadership to attend a conference that included agenda items designed for CTR members to gain a better appreciation of each other's work. Participants' conference feedback suggested that they felt one of the greatest benefits was gaining an appreciation for the challenges facing their colleagues in other cores (DaCCoTA, 2019). Thus, to continue to reinforce the culture of cooperation, the conference is now an annual event. Finally, participant feedback from our tabletop exercise indicated that the process of working together to streamline SOPs also facilitated relationship building.

Relevance. Just as the information source influences credibility and thus the likelihood of action, so does the perceived relevance of the feedback. Feedback mechanisms with too much noise make it difficult to extract the pertinent information (Bell et al., 2001; Buckley et al., 1994), potentially obscuring relevant feedback. All CTR

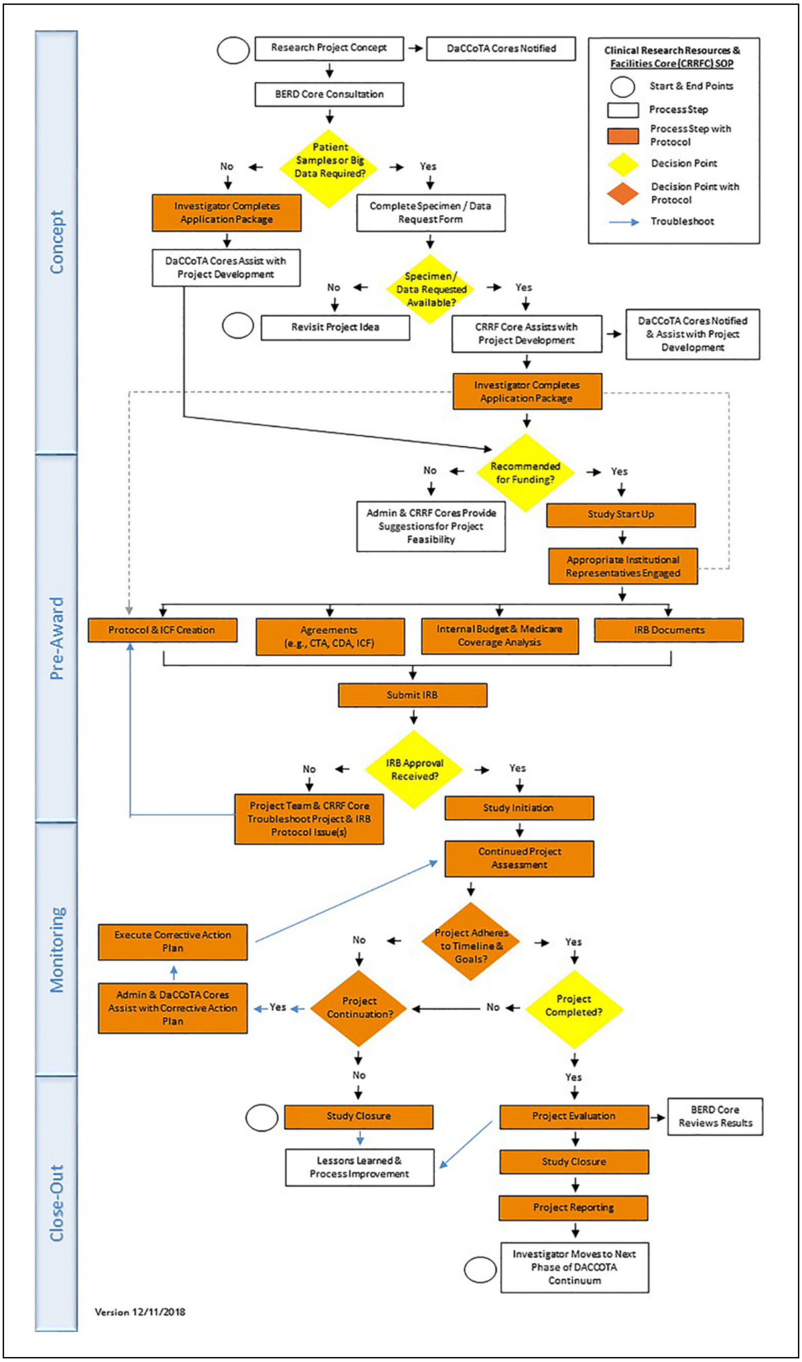

Figure 3. CTR core standard operating procedure.

investigators were mentored in understanding what constituted an appropriate core request. Limiting correspondence to relevant requests helped reduce noise and ensure timely feedback. Similarly, CTR staff were trained to reduce system noise by only using CTR feedback mechanisms for CTR support: using these feedback mechanisms to organize social events, sell merchandise and so forth was discouraged.

Specificity. Too much, or irrelevant, information can convolute feedback resulting in processing delays and even worse an inappropriate system response which will lead to further delays and system waste (Choi & Myung, 2017). As an illustration of how the application of the specificity principle was applied in our evaluation, we adopted a policy of sending recommendations sequentially, not simultaneously, to CTR administration for consideration. Leadership then had two weeks to close the feedback loop by informing us whether they chose to act on the recommendation. No additional recommendations were provided until the feedback loop was closed. The CTR director found this strategy helpful in facilitating decision-making.

Timeliness. The timeliness of feedback could significantly influence whether the CTR meets its goals. For example, CTR leadership was acutely aware of the need to ensure IRB approval before project data collection could proceed. However, leadership recognized that many novice investigators may be unfamiliar with the IRB process and were concerned whether they might inadvertently proceed without IRB approval. Therefore, training and mentoring emphasized the importance of waiting for IRB feedback and approval before proceeding.

Frequency. Feedback which is sent too frequently risks being viewed as system noise and ignored, even if the feedback is credible, sent in a timely fashion and is relevant. However, insufficiently frequent feedback may delay timely implementation of corrective actions. For example, the CTR invested in developing an IT infrastructure designed to flag potential communication breakdowns between cores. Our evaluation team is working with the IT staff to adjust the underlying algorithm such that it only provides feedback for recurring communication problems (i.e., two times in a row) between the same two system actors. This allows system actors an opportunity to self-correct and avoids overwhelming CTR staff and investigators with feedback.

Cascading failures. Because systems are interconnected, a problem occurring at one point in the system can be passed down to other parts of the system. This is also known as a system cascading failure or domino effect (Buldyrev et al., 2010; Renger, Foltysova, Ienuso, et al., 2017). Thus, when an inefficiency was discovered, we examined whether the problem was indeed occurring at the observed process step or whether it was due to an upstream dependency. For example, in Year 1, we learned that a system surge occurred because initial NIH funding was delayed by 3 months, and it took three additional months once funding had been achieved to create the relevant structures, hire staff, design and issue requests for proposals and evaluate proposals that were eventually submitted. The delayed funding placed pressure on our CTR to complete a year's work in a 6-month span. In turn, this shortened timeframe may not have offered interested investigators the sufficient time necessary to draft proposals. Thus, investigators may have hastily submitted proposals with avoidable flaws, while others may simply have chosen not to submit their projects for funding consideration. Understanding this cascading failure was caused by a funding nuance in Year 1 and that Year 2 funding was being provided on schedule at the start of the fiscal year, we did not recommend any corrective actions to the SOPs. Presumably, having a full 12 months will allow interested investigators multiple opportunities over the year to apply for funding.

Reflex arcs. Sometimes systems can bypass certain process steps in an effort to improve efficiency (Dewey, 1896). For example, to encourage collaboration, our CTR developed a portal where novice investigators can search the DaCCoTA network for senior investigators with shared interests. Through our structured interview process, we learned one of the CTR senior investigators requested removal from the pool because he was being inundated with collaboration requests from junior investigators. The solution to this problem was to allow senior investigators direct control by creating an activation/deactivation feature. This avoided the problem of relying on CTR leadership having to monitor and control access by placing control directly in the user's hands.

Lessons learned in evaluating CTR infrastructure efficiency

SET recommends following three steps in defining a system. However, in practice, we found it more useful to view the three steps as an iterative rather than sequential process. Specifically, we used findings from PFM (i.e., third step) to inform recommendations pertaining to adjusting system boundaries (i.e., first step) and relationships between cores (second step). The aforementioned reworking of the IRB process is an example of how relationships between cores were altered based on an evaluation of individual core processes.

Many of the core directors were appreciative of the PFM exercise, although they found it challenging to make explicit what they had done intuitively throughout their careers. Many also realized the immediate benefits of having SOPs, in which they could also serve as a training tool for core members to ensure standardization. The co-directors also came to realize the importance of micro- and macro-level training. The former refers to core staff understanding of internal core processes necessary for project-level support, while the latter refers to the need for core staff to understand how they fit in the larger CTR picture and how their work influences other cores and overall CTR success.

Engaging in the PFM exercise itself was a form of developmental evaluation (Patton, 1994). As directors and co-directors worked through the PFM exercise, they realized inefficiencies of their process steps and made real-time modifications. Thus, the validated SOPs represented new, improved operating procedures rather than a depiction of the current operating reality. Perhaps more rewarding from the CTR leadership perspective was core directors also gained a deeper appreciation of operating

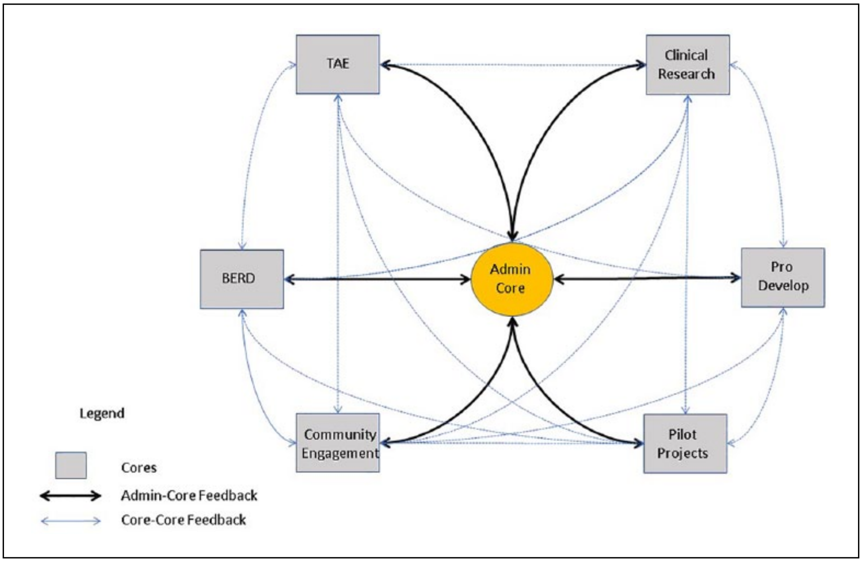

Figure 4. CTR system structure for evaluation purposes.

within the system. Core directors often noted how their processes would need to change because they were expected to engage with other cores to move a project through its completion and along the translational continuum. In short, the SOP mapping process reinforced the system concept among system elements.

We also learned evaluating CTR efficiency required a reconceptualization from an organizational structure to an operational structure. The reconceptualization was necessary to capture the reality in how cores were connected to each other and the responsibility of the administrative core for general system oversight (see Figure 4).

SET proved useful in providing us with direction in evaluating the efficiency of the CTR infrastructure. Analysing the SOPs from several perspectives helped detect different types of system inefficiencies. While there were some intra-core inefficiencies, the majority of inefficiencies were related to inter-core connectivity. Discovering continued connectivity-related inefficiencies reinforced and affirmed our use of SET as a better option for evaluating system efficiency than, for example, a programme evaluation approach that focuses on evaluating the fidelity with which core-specific protocols are implemented.

Once an inefficiency is identified, it is important to explore why it exists. SET suggests that system inefficiencies can often be attributed to one of the four factors: a lack of leadership, training failures, information technology failures and/or a poor organizational culture. These factors themselves are inextricably connected. For example, many CTR feedback loop failures were linked to IT problems. Upon deeper analysis, the IT failure could be attributed to a lack of staff training, failure of leadership to provide needed IT resources or the failure to appreciate the consequences to other CTR cores and the success of the CTR as a whole.

Strong and committed CTR and core leadership is essential to CTR efficiency. As such, the CTR director was mandated that all those in leadership positions attend a full-day symposium devoted to understanding the CTR goal and the interdependence of cores in meeting the goal.

Using SET led to useful evaluation findings. The utility of evaluation findings is an important standard in understanding the evaluation's value (Patton, 2008). To date, the TAE has made 23 recommendations of which 17 were accepted. Some of these were acted upon immediately, while others requiring broader group input were placed on the monthly standing meeting agenda.

Evaluating CTR infrastructure effectiveness

To meet this evaluation purpose, the NIH recommends developing logic models for each core (NIH, 2016). However, our concern in developing independent core logic models was twofold. First, we knew logic models would fail to capture the relationship between cores needed for the CTR to be effective in meeting its overarching goals. Second, we were concerned that an evaluation process pursing developing logic models might inadvertently reinforce a siloed rather than cooperative approach between cores. Our solution to this dilemma was to build on the work of Rogers (2008) by ensuring each core logic model was aligned to the same long-term goal. We used source documentation to develop an initial draft of each core logic model (Renger, 2010). The draft logic models were then shared with core leadership for validation. The core logic models were then depicted jointly as shown in Figure 5.

From Figure 5, it is evident that the long-term outcomes being collected are the same for all cores. Furthermore, the long-term outcome measures are typical success indicators in the research environment: the number, size and type of grant awards, publications, citations and so forth. From Figure 5, it is also evident that each core had unique immediate and intermediate outcomes needing to be evaluated. This would be expected as each core uniquely contributes to investigator independence.

Our next challenge was evaluating CTR effectiveness at moving project ideas along the translational continuum (CTR goal 2). Cancer research can take decades to move from the laboratory to the population as a whole (Simonds et al., 2013). This posed an evaluation challenge because the time needed to pass through the continuum would likely exceed the CTR funding duration. We are currently working with our evaluation counterparts at other CTRs to explore whether methods akin to tracing ancestry could be useful in evaluating this goal.

Lessons learned in evaluating CTR system effectiveness

From a systems perspective, the concept of developing core logic models to evaluate effectiveness seems counterintuitive. CTR cores are seen to rise and fall with the tide together as it were; the cores have a shared responsibility in the CTR meetings its overarching goals. Success cannot be solely attributed to one core as all the cores are intended to work together collaboratively and synergistically. Therefore, from a systems perspective, we used overall measures of effectiveness and did not attempt to partition success based on individual core contributions.

		Outcomes		
Program Assumptions	Strategy	Immediate	Intermediate	Long-term
Understanding how to engage community to create project buy-in is essential for project success.	Investigator completes cultural sensitivity training.	Investigator develops research proposal in concert with target population.	Investigator secures community cooperation.	Investigator secures independent funding
Understanding how to match the design and methods to the research question is essential for project success.	Investigator completes research training modules	Investigator demonstrates appreciation of mixed methods.	Proposal receives maximum design points.	
Junior investigators require guidance in the research process.	Mentors complete training courses.	Mentor identifies areas of weakness and structures learning opportunities to address weakness.	Self-efficacy of junior investigator to complete the proposal process is high.	

Community Engagement Core	
Research Design Core	
Professional Development Core	

Figure 5. Professional development core logic model.

Our use of multiple logic models aligned towards a common goal proved useful at showing how cores are working together even though they each contribute uniquely to investigator independence. Our solution was, of course, imperfect. Each core's logic model still depicts relationships as being linear, a result of the 'if-then' methodology used to derive the underlying operating assumptions (McLaughlin & Jordan, 1999; Pell Institute, 2019). Furthermore, and perhaps more importantly, our series of logic models still did not capture nonlinear relationships, such as feedback and communication loops necessary for core collaboration.

Conclusion

The focus of this manuscript was on sharing lessons learned in evaluating a research support infrastructure consisting of several cores (units) designed to work together collaboratively. The NIH guidance suggests using a logic model approach to evaluate the unique contributions of each core in meeting the CTR overarching goals. While logic models are useful in helping evaluate core impact, it was our judgement that they did not adequately evaluate a key CTR design feature; namely the need for cores to work together collaboratively to meet the overarching goals. Furthermore, it was our concern that only using logic models might inadvertently reinforce a siloed, rather than a collaborative approach. We therefore decided to complement the logic model approach with a system approach (i.e., SET).

We made significant upfront investments in applying SET to evaluate CTR infrastructure efficiency. We were rewarded by having SOPs that we could use as the standard of acceptability against which to evaluate system efficiency. What we did not anticipate was the challenge in evaluating the level of SOP detail. Evaluating every single SOP step was overwhelming and we worried about the extent to which attempting to do so would undermine the willingness of system actors to participate in the evaluation (Sanders, 1994; Taylor-Powell & Boyd, 2008). Borrowing from other disciplines, such as emergency preparedness, proved useful in allowing us focus on major SOP events and adhere to the feasibility evaluation standard (Sanders, 1994). We defined these major events as those where the system is most likely to fail, namely 'hand-off' points between cores and in closing feedback loops with quality information.

Overall, the systems approach to evaluating the CTR infrastructure proved useful as evidenced by the quick adoption of many evaluation recommendations. We realize other systems research and approaches from other disciplines may be useful in further augmenting how one defines and evaluates a system (e.g., Ackoff, 1971; Gharajedaghi, 2011; Midgely, 2007; Reynolds & Holwell, 2010). However, the challenge will be how to apply systems concepts from other disciplines in a comprehensive, utility-focused evaluation framework.

Using an SET proved useful in providing near real-time actionable recommendations. One of the greatest advantages of SET was it provided a richer and more meaningful implementation evaluation (i.e., focused on interconnections) than would have been possible using a programme evaluation approach grounded in logic models (i.e., focused on siloed core process). However, a systems evaluation approach is more costly to design and implement than a traditional programme evaluation approach using logic models. Nevertheless, as the saying goes 'you get out what you put in'. Our hope is that other similarly structured programmes might be able to reduce their evaluation costs by learning from our experiences.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship and/or publication of this article: The work is supported by Supported by NIH 1 U54 GM128729.

ORCID iD

Ralph Renger https://orcid.org/0000-0002-0224-1444

Note

1. Hereinafter, we use the terms 'our' and 'we' to represent the authors of this article and TAE staff members.

References

Ackoff, R. (1971). Towards a system of systems concepts. *Management Science*, *7*(11), 661–671. https://doi.org/10.1287/mnsc.17.11.661

American Evaluation Association. (2018, September 9). *Principles for effective use of systems thinking in evaluation: Systems in evaluation topical interest group*. https://www.systemsinevaluation.com/wp-content/uploads/2018/10/SETIG-Principles-FINAL-DRAFT-2018-9-9.pdf

Arnheiter, E. D., & Maleyeff, J. (2005). The integration of lean management and Six Sigma. *The TQM Magazine*, *17*(1), 5–18. https://doi.org/10.1108/09544780510573020

Banathy, B. H. (1992). *A systems view of education: Concepts and principles for effective practice*. Educational Technology.

Bell, P. A., Green, T., Fisher, J. D., & Baum, A. (2001). *Environmental psychology*. Oxford University Press.

Bentley, T. G., Effros, R. M., Palar, K., & Keeler, E. B. (2008). Waste in the U.S. health care system: A conceptual framework. *Milbank Quarterly*, *86*(4), 629–659. https://doi.org/10.1111/j.1468-0009.2008.00537.x

Birt, L., Scott, S., Cavers, D., Campbell, C., & Walter, F. (2016). Member checking: A tool to enhance trustworthiness or merely a nod to validation? *Qualitative Health Research*, *26*(13), 1802–1811. http://doi.org/10.1177/1049732316654870

Buckley, C., Salton, G., & Allan, J. (1994). The effect of adding relevance information in a relevance feedback environment. In B. W. Croft & C. J. van Rijsbergen (Eds.), *SIGIR'94* (pp. 292–300). Springer.

Buldyrev, S. V., Parshani, R., Paul, G., Stanley, H. E., & Havlin, S. (2010). Catastrophic cascade of failures in interdependent networks. *Nature*, *464*(15), 1025–1028. https://doi.org/10.1038/nature08932

Choi, N., & Myung, R. (2017). Feedback frequency effect on performance time in dynamic decision-making task. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *61*(1), 188–192. https://doi.org/10.1177/1541931213601531

Dakota Cancer Collaborative on Translational Research. (2019, June 8). Dakota Cancer Collaborative On Translational Activity. *Paper presented at the Inaugural Symposium. School of Medicine & Health Sciences*, Grand Forks, ND, United States.

Dewey, J. (1896). The reflex arc concept in psychology. *Psychological Review*, *3*(4), 357–370. https://doi.org/10.1037/h0070405

Gharajedaghi, J. (2011). *Systems thinking: Managing chaos and complexity: A platform for designing business architecture* (3rd ed.). Elsevier.

Goodwin, B., & Miller, K. (2012). Research says good feedback is targeted, specific, timely. *Educational Leadership*, *70*(1), 82–83.

Green, L. W., Kreuter, M. W., Deeds, S. G., Partridge, K. B., & Bartlett, E. (1980). *Health education planning: A diagnostic approach*. Mayfield Publishing.

Hoffart, N. (1991). A member check procedure to enhance rigor in naturalistic research. *Western Journal of Nursing Research*, *13*(4), 522–534. https://doi.org/10.1177/019394599101300408

Ilgen, D. R., Fisher, C. D., & Taylor, M. S. (1979). Consequences of individual feedback on behavior in organizations. *Journal of Applied Psychology*, *64*(4), 349–371. https://doi.org/10.1037/0021-9010.64.4.349

Institute for Clinical and Translational Research. (2019). What are the T0 to T4 *research classifications?* https://ictr.wisc.edu/what-are-the-t0-to-t4-research-classifications/

London, M., & Beatty, R. W. (1993). 360-degree feedback as a competitive advantage. *Human Resource Management*, *32*(2–3), 353–372. https://doi.org/10.1002/hrm.3930320211

Lub, V. (2015). Validity in qualitative evaluation: Linking purposes, paradigms, and perspectives. *International Journal of Qualitative Methods*, *14*(5), Article 621406. https://doi.org/10.1177/1609406915621406

McLaughlin, J. A., & Jordan, G. B. (1999). Logic models: A tool for telling your programs performance story. *Evaluation and Program Planning*, *22*(1), 65–72. https://doi.org/10.1016/S0149-7189(98)00042-1

Meadows, D. H., & Wright, D. (2008). *Thinking in systems: A primer*. Chelsea Green Publishing.

Midgely, G. (2007). *Systems thinking for evaluation: Systems concepts in evaluation: An expert anthology*. Edge Press.

Nave, D. (2002). How to compare six sigma, lean and the theory of constraints. *Quality Progress*, *35*(3), 73–80.

National Institutes of Health. (2016, July 11). *CTSA program common metric operational guideline: Careers in clinical and translational research*. https://www.tuftsctsi.org/wp-content/uploads/2016/07/Careers-in-CT-Research-Operational-Guidelines-7.11.16-v-1.2.pdf

National Institutes of Health. (2017). *Institutional Development Award (IDeA) program infrastructure for clinical and translational research (IDeA-CTR)* [PAR-17-304]. https://grants.nih.gov/grants/guide/pa-files/par-17-304.html

Patton, M. Q. (1994). Developmental evaluation. *Evaluation Practice*, *15*(3), 311–319. https://doi.org/10.1177/109821409401500312

Patton, M. Q. (2008). *Utilization-focused evaluation*. SAGE.

Pell Institute. (2019, June 22). *Evaluation toolkit; Using a logic model*. http://toolkit.pellinstitute.org/evaluation-guide/plan-budget/using-a-logic-model/

Renger, R. (2010). Constructing and verifying program theory using source documentation. *Canadian Journal of Program Evaluation*, *25*(1), 51–67.

Renger, R. (2015). System evaluation theory (SET): A practical framework for evaluators to meet the challenges of system evaluation. *Evaluation Journal of Australasia*, *15*(4), 16–28. https://doi.org/10.1177/1035719X1501500403

Renger, R. (2016). Illustrating the evaluation of system feedback mechanisms using system evaluation theory (SET). *Evaluation Journal of Australasia*, *16*(4), 14–20. https://doi.org/10.1177/1035719X1601600403

Renger, R., & Bourdeau, B. (2004). Strategies for values inquiry: An exploratory case study. *American Journal of Evaluation*, *25*(1), 39–49. https://doi.org/10.1177/109821400402500103

Renger, R., Foltysova, J., Ienuso, S., Renger, J., & Booze, W. (2017). Evaluating system cascading failures. *Evaluation Journal of Australasia*, *17*(2), 29–36. https://doi.org/10.1177/1035719X1701700205

Renger, R., Foltysova, J., Renger, J., & Booze, W. (2017). Defining systems to evaluate system efficiency and effectiveness. *Evaluation Journal of Australasia*, *17*(3), 4–13. https://doi.org/10.1177/1035719X1701700302

Renger, R., Keogh, B., & Hawkins, A. (2018). Reworks: A robust system efficiency measure. *Evaluation Journal of Australasia*, *18*(3), 183–191. https://doi.org/10.1177/1035719X18796611

Renger, R., McPherson, M., Kontz-Bartels, T., & Becker, K. (2016). Process flow mapping for systems improvement: Lessons learned. *Canadian Journal of Program Evaluation*, *31*(1), 109–121.

Renger, R., Renger, J., Renger, J. A., Donaldson, S., & Hart, G. (in press). Applying systems thinking concepts to evaluate systems. *Canadian Journal of Program Evaluation*.

Reynolds, M., & Holwell, S. (2010). *Systems approaches to managing change: A practical guide*. Springer.

Rogers, P. J. (2008). Using programme theory to evaluate complicated and complex aspects of interventions. *Evaluation*, *14*(1), 29–48. https://doi.org/10.1177/1356389007084674

Rubio, D. M., Schoenbaum, E. E., Lee, L. S., Schteingart, D. E., Marantz, P. R., Anderson, K. E., Platt, L. D., Baez, A., & Esposito, K. (2010). Defining translational research: Implications for training. *Academic Medicine*, *85*(3), 470–478. https://doi.org/10.1097/ACM.0b013e3181ccd618

Sanders, J. R. (1994). *The program evaluation standards: how to assess evaluations of educational programs*. SAGE.

Simonds, V. W., Wallerstein, N., Duran, B., & Villegas, M. (2013). Community-based participatory research: Its role in future cancer research and public health practice. *Preventing Chronic Disease*, *10*, Article 120205. https://doi.org/10.5888/pcd10.120205

Stazyk, E. C., & Davis, R. S. (2019). Transformational leaders: Bridging the gap between goal ambiguity and public value involvement. *Public Management Review*, *22*(3), 364–385. https://doi.org/10.1080/14719037.2019.1588357. https://www.tandfonline.com/doi/abs/10.1080/14719037.2019.1588357?journalCode=rpxm20

Taylor-Powell, E., & Boyd, H. H. (2008). Evaluation capacity building in complex organizations. *New Directions for Evaluation*, *120*, 55–69. https://doi.org/10.1002/ev.276

United States Department of Homeland Security. (2020). *Homeland Security Exercise and Evaluation Program (HSEEP)*. https://www.fema.gov/media-library-data/1582316517342-782e4e95508f4a63cf2e62ae2b5f5f55/Homeland-Security-Exercise-and-Evaluation-Program-Doctrine-2020-Revision-2-2-21-508.pdf

United States Office of Personnel Management. (2019, June 23). *Feedback is critical to improving performance*. https://www.opm.gov/policy-data-oversight/performance-management/performance-management-cycle/monitoring/feedback-is-critical-to-improving-performance/

Williams, B. (2010, November 12). Fitting the key to the lock: Matching systems methods to evaluation questions. *Paper presented at the 24th Annual Conference of the American Evaluation Association*, San Antonio, TX, United States.

Author biographies

Ralph Renger, PhD, MEP specializes in developing new evaluation methods such as mainstreaming of logic modeling approaches and more recently systems evaluation theory (SET). After serving for 25 years as a professor at the University of Arizona and the University of North Dakota he now serves as president of Just Evaluation Services, LLC

Marc D Basson, MD, PhD, MBA is Senior Associate Dean for Medicine and Research at the School of Medicine and Health Sciences of the University of North Dakota and Principal Investigator of the DaCCoTA consortium.

Gary Hart, PhD is the Director for the Center for Rural Health at the School of Medicine and Health Sciences of the University of North Dakota. He is also the co-director of the Tracking and Evaluation Core of the DaCCoTA consortium.

Rick Van Eck, PhD is the Associate Dean for Teaching and Learning and the founding Dr. David and Lola Rognlie Monson Endowed Professor in Medical Education at the University of North Dakota School of Medicine and Health Sciences. He is also the co-director of the Tracking and Evaluation Core of the DaCCoTA consortium.

Eric Souvannasacd is a Systems Evaluator at the Dakota Cancer Collaborative on Translational Activity at the University of North Dakota School of Medicine and Health Sciences. He is pursuing a PhD in Educational Foundations and Research at the College of Education and Human Development at the University of North Dakota.

Jessica Renger is a current graduate student in the Evaluation and Applied Research Methods PhD program at Claremont Graduate University. Before coming to Claremont, she received a bachelor's degree in Neuroscience from the University of Arizona. As a practitioner, she has focused on evaluating health and education programs in Arizona, California, North Dakota, and Wyoming.

Jirina Renger, DrPH graduate from University of Arizona, College of Public Health. She currently serves as a core faculty member in the DrPH and PhD Public Health programs in the College of Health Sciences at Walden University. Her area of specialization is theory-driven evaluation and program theory reconstruction.