



First steps: creating an initial program theory for a realist evaluation of Healthy Start-Départ Santé intervention in childcare centres

Fiona Fick & Nazeem Muhajarine

To cite this article: Fiona Fick & Nazeem Muhajarine (2019): First steps: creating an initial program theory for a realist evaluation of Healthy Start-Départ Santé intervention in childcare centres, International Journal of Social Research Methodology, DOI: [10.1080/13645579.2019.1595375](https://doi.org/10.1080/13645579.2019.1595375)

To link to this article: <https://doi.org/10.1080/13645579.2019.1595375>



Published online: 25 Mar 2019.



Submit your article to this journal [↗](#)





Article views: 51



View Crossmark data [↗](#)



First steps: creating an initial program theory for a realist evaluation of Healthy Start-Départ Santé intervention in childcare centres

Fiona Fick  and Nazeem Muhajarine 

Department of Community Health and Epidemiology, College of Medicine, University of Saskatchewan, Saskatoon, Saskatchewan, Canada

ABSTRACT

Realist evaluation (RE) is increasingly adopted to assess the effectiveness of social programs. RE goes beyond measurement of outcomes, providing insight into how and why programs work by focusing on how participants experience and respond to programs within certain contexts. A RE was conducted on Healthy Start-Départ Santé (HSDS), an early childhood development program in Canada, created to address the prevalence of overweight and obese children. The first step of the RE was to create an initial program theory (IPT) to identify key contexts and mechanisms relating to the success of the program. Realist methodology can be challenging, time consuming, and resource intensive and there are few practical examples of how to conduct a RE. This article aims to add to the realist methodological literature by describing the construction of the IPT, reflecting on the process used, and providing a starting point for future realist practitioners.

ARTICLE HISTORY

Received 12 October 2018
Accepted 12 March 2019

KEYWORDS

Realist evaluation; realist methodology; program theory; childhood development; childcare center

Introduction

All truths are easy to understand once they are discovered; the point is to discover them. Galileo Galilei

The most effective way to do it, is to do it – Amelia Earhart

The use of realist methods to evaluate social programs is increasing (Salter & Kothari, 2014), spurring the need for clarity in realist methodology. Reviews of realist evaluations (RE) show that researchers find realist methodology challenging, time-consuming, and resource intensive and often fail to comply with key tenants of RE. For example, it has been observed that some researchers use the Context-Mechanism-Outcome configuration as a descriptive tool rather than to draw causal connections between the context and mechanisms of the intervention and the outcomes as well as incorrectly distinguishing between contexts and mechanisms (Coryn, Noakes, Westine, & Schröter, 2011; Marchal, Van Belle, Van Olmen, Hoérée, & Kegels, 2012; Salter & Kothari, 2014). Despite the challenges of realist methodology, the use of theory-based approaches to evaluation of programs are acknowledged to provide a deeper understanding of how, when, and for whom a program works (Donaldson & Lipsey, 2006; Rogers, 2008).

At its core, a RE consists of constructing an initial program theory, testing the theory using suitable methods (i.e. quantitative, qualitative, or mixed), and refining the initial theory to create a middle-range theory that can be applied to different contexts. This is an iterative rather than

linear process as the theory is repeatedly tested and refined with accumulating knowledge. Although helpful resources for assessing quality and standards for reporting have recently been created (Wong et al., 2017), there are few practical examples of how to conduct REs or how to construct a program theory (Linsley, Howard, & Owen, 2015; Shearn, Allmark, Piercy, & Hirst, 2017; Van Belle, Marchal, Dubourg, & Kegels, 2010). This article aims to add to the realist methodological literature by describing in detail the steps taken to create an initial program theory related to the realist evaluation of a population base intervention focusing on preschool age children: the Healthy Start-Départ Santé program.

Before discussing the specifics of RE, it is worthwhile to provide a short overview of the underlying philosophy. Pawson and Tilley (1997), in their seminal publication, draw upon scientific realism as their philosophical base. Realism is located between positivist and constructivist accounts of science, and acknowledges that although all knowledge is incomplete and fallible (Maxwell, 2012), there exists a reality independent of our beliefs and constructions that is discoverable through empirical inquiry (Sayer, 1992). However, this reality is filtered through human volitions, culture, and language, which produce constant change within social institutions (Greenhalgh, Wong, Westhorp, & Pawson, 2011). Scientific knowledge advances through positing theories of causation, testing, revising and retesting of these theories, based on phenomenon that are unobservable yet real. These theories provide an organizing framework that extracts essential conditions for program success in different contexts (Pawson & Tilley, 1997). Knowledge gained through testing of these theories in different contexts provides information that, in turn, could be applied in subsequent times and circumstances to achieve desirable outcomes.

Pawson and Tilley (2004, p. 3) believe that 'programs are theories incarnate.' All programs begin as ideas about how to bring about change in a particular area or as hypotheses about social improvement. However, in our experience, program theories are often implicit as those who create programs seldom explicate how and why they expect a program to work. A theory can be defined as, 'a set of interrelated constructs, definitions, and propositions that present a systematic view of phenomena by specifying relations among variables, with the purpose of explaining and predicting phenomena' (Kerlinger, 1986, p. 9). Program theories are central to realist evaluation as they provide plausible explanations as to how and why a program works (or does not work) within a particular context. Realist evaluations are theory-driven, with the unit of analysis being the program theory rather than program activities. This changes the focus of the evaluation from the success of outcomes to the causal theory underpinning the program (Pawson & Tilley, 1997) and moves the evaluation from the question of 'does it work' to the more complex and useful question of 'what works, for whom, in what contexts, and how' (Pawson, 2006). This provides a deeper level of understanding and often generates more useful information for decision-makers. Successful replication of the program involves the transfer of the program theory and not merely the program activities. Thus, the first step in a RE is to create an initial program theory (IPT; Pawson & Tilley, 2004; Wong et al., 2017). The ultimate aim of a RE is to create a transferrable middle-range theory about how outcomes are produced in different contexts.

A commonly used definition of middle range theories comes from Merton (1968, p. 68), who describes it as:

Limited sets of assumptions from which specific hypotheses are logically derived and confirmed by empirical investigation. These theories do not remain separate but are consolidated into wider networks of theory.. [that are] sufficiently abstract to deal with different spheres of social behavior and social structure... The middle range orientation involves the specification of ignorance. Rather than pretend to knowledge where in fact it is absent, it expressly recognizes what still must be learned in order to lay the foundations for still more knowledge.

Pawson (2010) elaborates that sufficient abstraction is a necessary tool in theory building. Concepts should not be tied so closely to a particular case that they cannot be transferred to others, nor should they be so abstract that they cover a multitude of cases and thus are incapable

of resonating with specificity with any one of them. Sufficient abstraction allows a specific case to fit within a broader explanatory schema. Middle-range theories should be able to bridge different spheres and draw together different fields of social behavior. However, it is not possible to specify a formula for an exact level of abstraction to be attained, as this would be determined by a multitude of factors. Furthermore, adaptive, cumulative explanations recognize that middle-range theory may sometimes be wrong. Social phenomena are rarely stable and will often change over time and place. Failure to explain anomalies will drive further inquiry, leading to an elaboration and cumulation of conceptual schemes.

Evaluations often focus on measuring outcomes in order to assess program effectiveness. However, ascertaining that a program had an effect does not explain how or why the effect occurred. Pawson (2006) points out that ‘what works?’ is a causal question that asks how social programs cause their effects. This is answered through the key realist idea of generative causation, which postulates that it is not the program itself that causes change, it is the reasoning and response of the participant to the program resources within a particular context that activate certain mechanisms that create change (Pawson & Tilley, 1997). Similarly, Weiss (1997, p. 46) states that ‘the mechanism of change is not the program service per se but the response that the activities generate.’ The base on which realist explanation rests is that ‘causal outcomes follow from mechanisms acting in contexts’ (Pawson & Tilley, 1997, p. 58), in other words, RE assumes that successful outcomes occur only when the program activates certain mechanisms in certain contexts. Thus, RE aims to understand the underlying mechanisms (M) which produce change, the contextual factors (C) necessary to activate these mechanisms, and how the combination of context and mechanisms produces outcomes (O). Pawson and Tilley describe this process as creating context-mechanism-outcome (CMO) configurations. These CMO configurations enable the evaluator to understand the inter-relationships between contexts, mechanisms, and outcomes, and thereby gain a deeper understanding of the success (or failure) of an intervention.

Social programs operate within a complex and messy social reality, which makes isolating and manipulating all potential explanatory variables extremely unlikely (Pawson, 2006). Rather than aiming to reduce complexity by treating contextual factors as confounding variables, realism incorporates complexity by emphasizing an understanding of how context shapes and affects causal mechanisms (Pawson & Tilley, 2004), making this a suitable philosophy to underpin evaluations of social programs existing within a social reality. RE can be particularly valuable when evaluating social interventions and gaining an understanding of the motivations and intentions of people vital to the success of the intervention (Pawson & Tilley, 1997).

Overview of Healthy Start-Départ Santé

Healthy Start-Départ Santé (HSDS) is a multi-level intervention created to address the prevalence of overweight and obese children in Canada. Canadian data indicate that 11.7% of children between 2–5 years of age are overweight and 7.3% are obese (Statistics Canada, 2015). Several studies show an association between obesity and sedentary patterns of behavior (e.g. Hills, King, & Armstrong, 2007) and unhealthy eating patterns (e.g. Bowman, Gortmaker, Ebbeling, Pereira, & Ludwig, 2004). HSDS aims to implement an evidence-based intervention designed to increase levels of physical activity and healthy eating in children attending licensed early learning childcare centers (ELCC). The goal of the intervention is to change the physical and social environment of ELCCs in order to increase opportunities for healthy eating and physical activity for children.

HSDS consists of a training session open to all staff members of ELCCs, a follow-up booster training session, and regular contact between HSDS and ELCC staff for a period of 10 months. A kit containing instructional materials and activity-enhancing tools is given to ELCC staff after the training session. A more detailed description of the program can be found in the protocol of the randomized controlled trial (RCT) assessing the efficacy of the intervention (Bélanger et al., 2016).

Although previous studies have found some evidence for the effectiveness of different interventions targeting physical activity and healthy nutrition in early childhood (Gordon, Tucker, Burke, & Carron, 2013; Temple & Robinson, 2014), the causal mechanisms underlying such interventions and how contextual factors influence the link between these causal mechanisms and outcomes remains unknown. This RE aims to develop a middle-range theory that explains how HSDS works to improve levels of physical activity and healthy nutrition in early childhood in order to understand key contexts and mechanisms relating to the success of the program. This will assist in determining how the future scale-up phase of this program can be implemented more effectively and have greater impact on the health and development of young children. The research question guiding this study is: How, for whom, and in what contexts does the HSDS intervention work to improve healthy eating behaviors and increase levels of physical activity in children in Early Learning Childcare Centers?

Ethics approval was obtained from the University of Saskatchewan Research Ethics Board (# 14–291).

Developing the initial program theory

Data collection and analysis

Multiple sources of data were used to create the Initial Program Theory (IPT): 1) a review of the internal HSDS literature and attendance at various HSDS meetings, 2) interviews with the program developers and trainers, and 3) peer reviewed literature on similar interventions targeting healthy nutrition and/or physical activity in ELCCs.

The first task was to become familiar with the HSDS program by interacting with the program stakeholders and reviewing the program literature. Different contexts (C), mechanisms (M), and outcomes (O) were drawn from the internal program literature and entered into a table. These were not necessarily related but provided a starting point for the coding of C, M, and Os. The two groups of people we felt could provide the best information about the program were the program developers and the HSDS trainers. First, we conducted a focus group and follow-up individual interviews with stakeholders involved in the creation and continuing improvement of the program (i.e. kinesiologists, health promotion specialists, epidemiologists, and nutritional experts). Five people attended the focus group and three program developers who could not attend the group session were interviewed individually. To begin, we explained how a realist evaluation worked and how it differed from other forms of evaluation. The concepts of context, mechanisms, and outcomes were also defined and explained. A brainstorming session, beginning with outcomes and ending with mechanisms resulted in a list of each.

Second, we conducted telephone interviews with four HSDS trainers who administer the training sessions, the booster sessions, and maintain contact with the ELCCs. The interviews with the trainers were slightly different to those with the program developers, as the trainers were not specifically asked to identify different contexts, mechanisms, and outcomes. Instead, they were asked about their experiences with the ELCC staff, for example, ‘Did the training seem to work better with a particular center or a particular type of center? Why do you think that is?’ We felt that a narrative account would give a greater depth of information from people who were intimately acquainted with the inner workings of the program, rather than creating lists of different contexts, mechanisms, and outcomes. Indeed, this type of questioning drew rich information that could then be coded into different C, M, and Os. However, establishing linkages between contexts, mechanisms, and outcomes was felt to be premature at this stage.

The C-M-O table created from the internal HSDS literature, the transcripts from the interviews, and the list of different contexts, mechanisms, and outcomes from the focus group were loaded for coding into NVivo 11 (QSR International, 2018), a qualitative data analysis software tool. Coding was specifically related to each context, mechanism, and outcome (e.g. ‘C – lack of equipment for physical activity’, ‘M – understanding of importance of physical activity’, ‘O – centers adapt their

environment'), but no connections or linkages were made between them at this stage. As some codes were related in meaning to others, they were amalgamated under an encompassing label. This is somewhat similar to thematic coding (Braun & Clarke, 2006); however, it was more of an exercise in amalgamation and categorization. Once codes had been grouped as far as possible, the list of codes was transferred to a Microsoft Word document to allow for greater flexibility.

The ultimate outcomes of the program, which are an increase in physical activity and increased healthy eating behaviors in children, were used to guide the next stage of analysis. A further C-M-O table was created, grouping the coded contexts, mechanisms, and outcomes against the desired outcomes of physical activity and/or healthy eating. Some codes were specific to one outcome and some could apply to both. Once these were listed in full, the contexts were separated into Internal and External Context in relation to the site of the program, the mechanisms were separated into Positive and Negative Mechanisms, and the outcomes were grouped according to three levels: Centre, Educators, and Children. The External Context included different geographical, cultural, and economic settings and the Internal Context included organizational issues such as funding, space, and equipment. Positive mechanisms included codes such as, 'Educators feel that change is achievable' and negative mechanisms included codes such as, 'Educator resistance to change'. Examples of outcome codes for different levels included, 'Centers adapt their environment', 'Educators facilitate physical activity', and 'Improved motor skills for children'.

The process of grouping and categorizing the initial codes took several iterations. For example, the initial codes of, 'M – Educators believe in the program', 'M – Educators understand why the program is important', and 'M – Educators understand benefits' were grouped and re-labelled as 'Staff understand the potential benefits of the program and believe the program is important to the healthy development of children.' Once all the initial codes had been grouped and renamed, the resulting sentences were further amalgamated and renamed as far as possible. At this point, connections were made between different contexts, mechanisms, and outcomes, resulting in 28 initial CMO configurations. Again, these CMO configurations were amalgamated where possible, separating them into categories of 'Successful Outcomes' and 'Unsuccessful Outcomes'.

On closer inspection, the data gathered thus far related more to the initial implementation of the program than the on-going effectiveness of the program. Upon reflection, this was not surprising as the sources of data so far were the internal HSDS literature documenting the creation of the program as well as implementation in various ELCCs, and the perspectives of the program developers and the trainers. However, in order to fully answer the evaluation questions, we needed to address the issue of on-going program effectiveness. For this, we turned to the peer-reviewed literature on similar interventions addressing physical activity and/or healthy nutrition in ELCCs. Most articles examined either implementation or effectiveness of programs with very few relating to both. As a result, we felt it would be more productive to analyze these separately with a view to potentially combining them later. Keyword searches rendered many results for the limited time frame, so we began with published literature reviews. Reading these reviews gave us information about relevant interventions, which we then mined for data regarding C, M, and Os. Because this was not a systematic review of different interventions, we are aware that some important information may have been missed. However, this was a pragmatic decision with the understanding that this material contributes towards an initial theory that will be confirmed, refuted, or refined by the data provided by the program users.

Again, two tables of C-M-Os were created (one for physical activity and one for nutrition interventions) with excerpts from journal articles copied and pasted into the relevant Context, Mechanism, or Outcome boxes. These were simply lists at this point and were not linked as CMO configurations. These tables were uploaded into the NVivo 11 software and coded using the same process as before. Once coded, the list of codes was copied into MS Word and categorized, grouped, and renamed as far as possible using the same process as before. Fifteen CMO configurations were formed from these data.

Although the categorization and renaming process had brought the data into various CMO configurations, we still had a large, fragmented collection of data. Finding a way to bring order to this was a challenge that was resolved one step at a time. First, instead of trying to combine the CMO configurations related to the implementation process and effectiveness, we decided to split the evaluation into Implementation and Effectiveness. This meant that we could see the data related to each more plainly. However, the data were still too noisy and did not yet give a clear picture of either of these. We further split Implementation into ‘Training’ and ‘Implementation Process’, which gave a clearer picture of how each of these could work. Effectiveness was split into ‘Staff Effectiveness’ and ‘Child Effectiveness’. Although the children are the target population, they will neither receive nor benefit from the program if the staff do not implement and continue with it. Again, this resulted in a clearer picture of how the program could work for both of these groups. We tried the effect of integrating both healthy nutrition and physical activity into the program theory for Child Effectiveness; however, it became overly complicated. Splitting it into each category gave a more coherent theory for both. Once we had theories for each stage, in a plausible causal process, we could see how they led to the next, i.e. Training → Implementation → Staff Effectiveness → Child effectiveness. This gave a clear, overarching framework for the IPT.

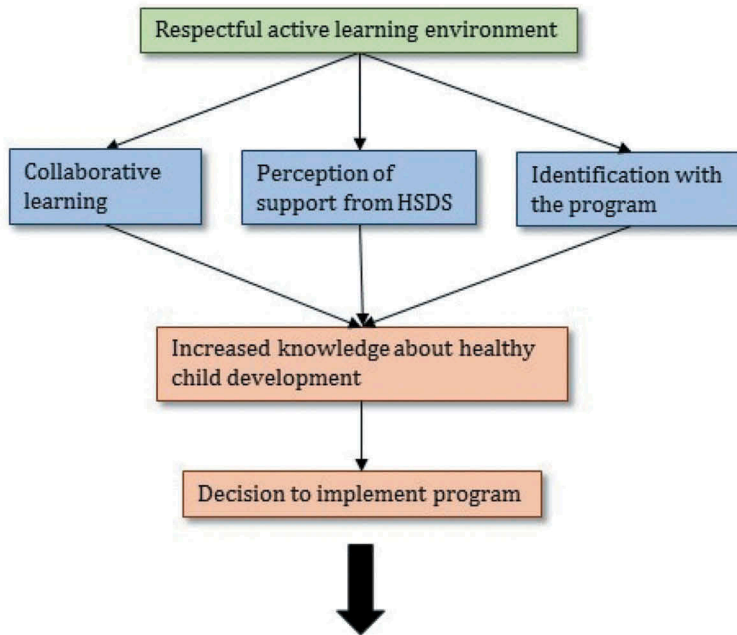
Initial program theory

The final IPT consisted of the following CMO configurations (see [Figure 1](#)):

- (1) **ELCC Staff Training:** Within a respectful active learning environment (C), participants feel supported by the HSDS trainers, learn from each other, and identify with a program that aligns with their goals and values (M). Participants become more knowledgeable about healthy childhood development and decide to implement the program (O).
- (2) **Program Implementation:** If ELCC staff have sufficient knowledge of healthy living for children, access to the necessary organizational resources for change, and a positive attitude and commitment towards facilitating healthy living in children (C), then they will carry out the implementation of the program (O) because they feel supported by the HSDS coordinators and other ELCC staff, they identify with the program, they feel a sense of responsibility for the healthy development of children, and they have gained a feeling of self-efficacy (M).
- (3) **Effectiveness (ELCC Staff):** If the early years organization has access to sufficient resources and the staff have a feeling of responsibility for healthy child development as well as sufficient knowledge and a positive attitude towards healthy living and program goals (C), then staff will role model and teach healthy behaviors and work towards improving healthy eating behaviors and increasing levels of physical activity in children (O) because they feel supported and able to make continual improvements to the program, they develop a sense of ownership of the program, and commit to active participation in the program (M).
- (4) **Effectiveness (Child):**
Physical Activity: If the organization provides opportunities for physical activity, the children have developed fundamental motor skills, and the group structure and dynamics facilitate participation (C), then children’s levels of autonomy will increase, leading to an increase in levels of physical activity (O) because children feel supported and encouraged by the staff, they have fun, and they feel a sense of self-efficacy (M).
Nutrition: If the organization has access to sufficient resources to supply healthy food/meals, the children have learned about healthy nutrition, and the group structure and dynamics facilitate consumption of healthy foods/meals (C), then their levels of autonomy will increase and they will have healthier eating behaviors (O) because children feel supported and encouraged by the staff, they feel a sense of self-efficacy, and they have positive sensory perceptions of food (M).

Each CMO configuration covers large concepts that require further explanation. However, including the explanation of each concept into the CMO resulted in a complex paragraph that was difficult to follow. To circumvent this challenge, we created a table of definitions to accompany the program theory, which included the definition of the concept as well as how the concept could facilitate or hinder the program (see [Table 1](#) for an example).

CMO 1: ELCC Staff Training



CMO 2: Program Implementation

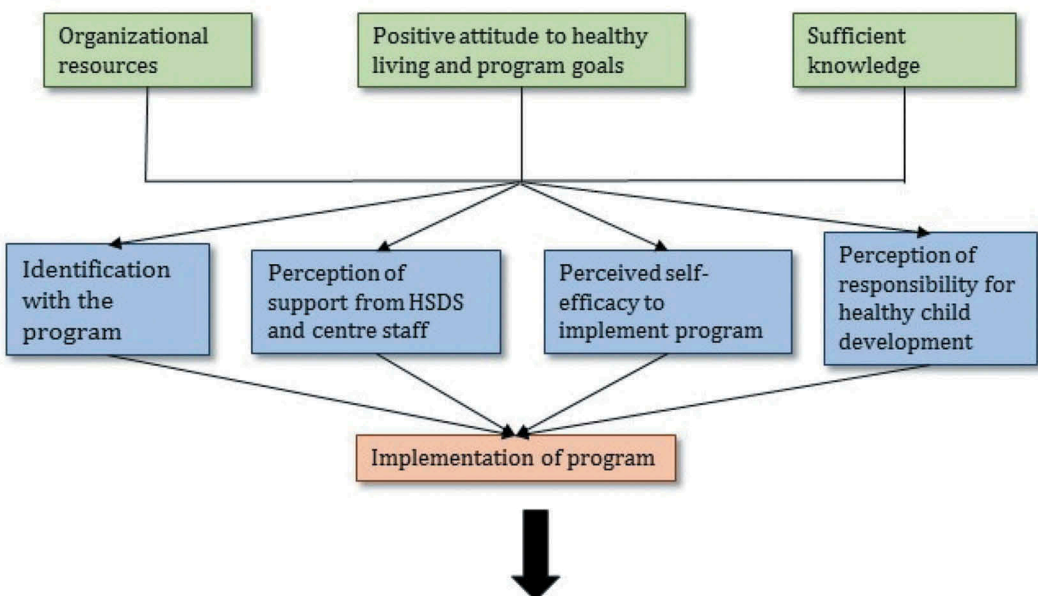
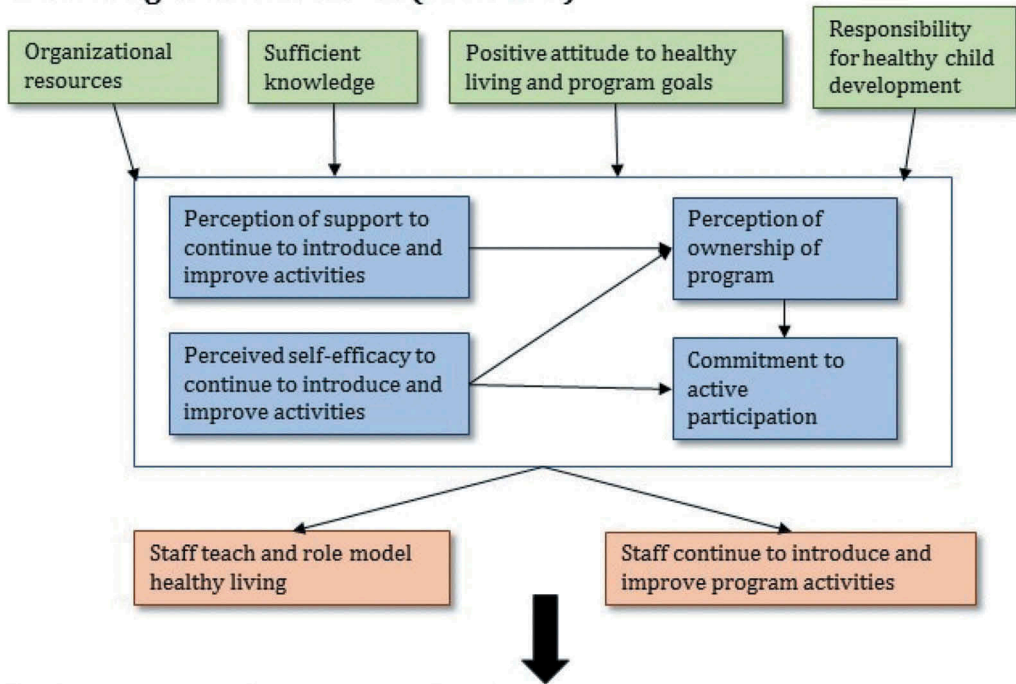


Figure 1. Healthy Start-Départ Santé Initial Program Theory.

CMO 3: Program Effectiveness (ELCC Staff)



CMO 4: Program Effectiveness (Child)

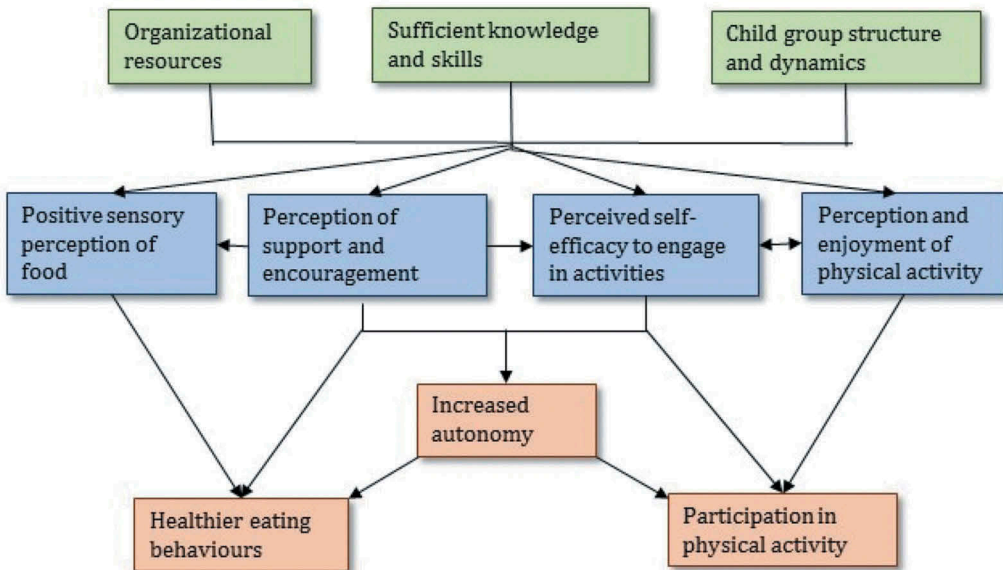


Figure 1. (Continued).

Next steps

In order to refine the IPT and create a middle-range program theory for HSDS, interviews will be conducted with Directors and Educators of ELCCs. These interviews will elicit confirmation or refutation of elements of the IPT through asking questions either directly related to the theory

Table 1. Example of Construct Definition.

	Label	Description
Mechanism	Self-efficacy	<p>Feeling of competence and ability to overcome challenges and make changes necessary to implement the program.</p> <p><i>Facilitates (High levels):</i> Staff feel able to create change and successfully implement the program activities. Staff feel able to participate in the activities and adapt them as necessary. Self-efficacy increases when staff see positive results in children.</p> <p><i>Hinders (Low levels):</i> Staff feel unable to overcome the barriers to change and implement the program activities. Staff feel they lack the competence to initiate and participate in different activities. Staff feel the proposed changes are not feasible for their environment.</p>

(e.g. Does the program line up with your personal values and goals around healthy child development?) or asking for their opinion on other aspects of the theory (e.g. Some people feel that centers have a responsibility to enhance healthy child development, whereas others feel that this is the responsibility of the parents. What are your thoughts on this?). The content of questions may change during the data collection process as iterative analysis takes place. If the quality of the data collected through the initial interviews is judged as insufficient to test the IPT, the focus of questions will be revised for further interviews.

A further step will be to underpin this with substantive social science theory. Substantive theories are established formal theories within a particular discipline that describe the process of reasoning or causation of actions. Substantive theories can identify rules that guide the activation of mechanisms, which in turn, can provide guidance for how to activate different mechanisms and the resulting outcomes. In order for behavior change to occur, pre-existing behaviors need to be adapted or new behaviors need to be established. We considered various behavior change theories and found that the Normalization Process Theory (NPT; May & Finch, 2009) gave insight into the Implementation component as it explains the process by which an intervention becomes embedded within an organization. We further found that the Self-Determination Theory (SDT; Deci & Ryan, 2008; Ryan & Deci, 2000), which explains the processes of intrinsic and extrinsic motivation, appeared to be a good fit with the Effectiveness component. These two theories will be utilized during the process of refinement and creation of the final program theory.

Reflections on the process of developing the initial program theory

The primary difficulty encountered in creating this IPT was lack of practical guidance. Although the steps of a RE are outlined in various articles and books, the ‘how to’ is not explicated. This is not an oversight; Pawson and Tilley (2004) see realist inquiry as a ‘logic of inquiry’ (p.10) rather than a research technique, or a method, and wanted to avoid prescriptions or a rigid structure in order to allow researchers/evaluators to use their own specific methods within this, that work best for them and for their project. Nevertheless, this does create ambiguity and uncertainty, particularly for those new to the field. A wide range of published practical examples may help those conducting RE to use what works for them in order to conduct rigorous evaluations in a timely manner.

One challenge in creating the IPT was dealing with a large amount of data and finding a balance between presenting this with enough detail to reflect the program accurately yet not so much detail that would render it confusing. As previously mentioned, we found the process of creating a table that reflected the definition of the construct as well as the way that construct facilitated or hindered the program to be extremely helpful in understanding how the program worked as well as why it might not work. Visually representing the IPT was a further challenge. Although we used boxes and arrows, there are alternative ways to depict these logical relationships and some may be more compelling than our visual diagrams.

We found it difficult to incorporate substantive theories while creating the IPT, as we did not yet know the main focus of the IPT. Once we identified this, we could then search for theories that would help to explain the processes in question. Thus, although some researchers advise that substantive theories should be utilized to create IPTs, we did not find this to be practical, particularly as HSDS had not looked to social science theories to underpin their proposed activities. However, substantive theories will be useful to underpin the final program theory.

One could ask whether a different person would have formulated the same IPT, which raises the topic of bias and subjectivity in the process. This is a valid question and one that should be reflected on by the researcher throughout the process. A bigger question though, is whether the IPT reflects the program accurately. This will be answered during the next phase of the RE, in which the IPT will be tested and refined through data from the program participants. Although we expect a certain level of revisions, refutation of the majority of the IPT would indicate a serious problem with the process or the data collected. Conversely, agreement with the majority of the IPT would indicate a good level of understanding of how the program works, regardless of whether another researcher would have created a different version of the IPT. Knowledge creation is a cumulative process, whereby new knowledge is built on previous knowledge. As such, realism does not claim final knowledge or an absolute certainty.

Realism is unique in its aim to uncover how and why an intervention works, rather than assessing issues such implementation fidelity or program effectiveness. A realist evaluation was proposed for HSDS to answer the puzzling question of why the program worked in some places but not in others, especially when the program activities and design were the same. What were the differences between ELCCs who succeeded in implementing and maintaining HSDS, yet often faced the same challenges as those who did not succeed in this? The strength of realist evaluation is the ability to answer these questions where many other types of evaluation cannot. However, uncovering these answers takes a great deal of time and effort and, like all research, has to be conducted transparently and rigorously. Additionally, similar to any evaluation, stakeholder engagement and program documentation contribute greatly to success.

Conclusions

As realist researchers and evaluators have indicated (Coryn et al., 2011; Marchal et al., 2012; Salter & Kothari, 2014), realist evaluations can be challenging to implement as there are few practical examples to guide the process. Pawson and Tilley (2004) themselves acknowledge that using a realist approach 'requires sustained thinking and imagination' to formulate program theories and design an evaluation to test them (p. 22). Nonetheless, realist methodology can produce innovative and insightful results different to those from other types of evaluation. This has been valued by the stakeholders of HSDS, particularly the ability to identify reasons for participant engagement (or lack thereof) as well as understanding the underlying factors not included in the program logic model. Indeed, one stakeholder expressed regret that she had not learned about this methodology earlier in her career.

This description of the process is offered as one example of creating an IPT; however, other realist practitioners may have used a different process. Realism is a methodology, not a method, and as such is an abstract concept. Concrete examples can aid in understanding abstract principles and can help to clarify how one might create a method useful to a different realist evaluation. The methods used in this process were literature reviews, focus groups, and interviews. We have outlined how we used these methods, but we are not expecting others to use these in exactly the same way. Certain steps used here will not be relevant to other evaluations, but we hope this can provide a starting point and help to highlight areas of congruency and difference. Conducting a RE requires extensive reading and, where possible, discussions with other realist researchers. Fortunately, the realist community is a generous one, willing to share their knowledge and experiences with others.

Acknowledgments

Authors are grateful for the important contribution of the program developers, the project manager and staff of Healthy Start-Départ Santé, as well as the support of Réseau Santé en Français de la Saskatchewan - Saskatchewan Network for Health Services in French. We especially acknowledge the input of the core Healthy Start-Départ Santé research team members. We also thank the childcare directors and educators for their valued contribution.

Authors' contributions

FF wrote the first draft of the manuscript. FF and NM designed the study and analyzed the data. All authors contributed to the manuscript and confirmed the final version.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the Public Health Agency of Canada under Grant 6282-15-2010/3381056-RSFS, the Consortium national de formation en santé under Grant 2014-CFME-01, the Heart and Stroke Foundation of Canada under Grant 2015-PLNI, and the Community Initiative Fund, Saskatchewan under Grant 14:949.

Notes on contributors

Fiona Fick received her Master of Arts in Social Psychology, researching burnout and work engagement in Long-Term Care Homes. She has focused on the evaluation of social programs, including employment training for newcomers as well as health interventions in long-term care and lower-income populations.

Nazeem Muhajarine, PhD, is a social epidemiologist and has focused on addressing widespread social issues, such as the health effects of income-based social inequality as well as reducing maternal mortality and improving infant health. He has been involved – often as the Principal Investigator – on more than 80 grants and contracts throughout his career and has published more than 170 papers in refereed journals, penned 13 peer-reviewed books and chapters in books and has written more than 70 technical reports for a wide range of agencies.

ORCID

Fiona Fick  <http://orcid.org/0000-0002-5537-9834>

Nazeem Muhajarine  <http://orcid.org/0000-0001-6781-5421>

References

- Bélanger, M., Humbert, L., Vatanparast, H., Ward, S., Muhajarine, N., Chow, A. F., ... Leis, A. (2016). A multilevel intervention to increase physical activity and improve healthy eating and physical literacy among young children (ages 3–5) attending early childcare centers: The Healthy Start-Départ Santé cluster randomized controlled trial study protocol. *BMC Public Health*, *16*(1), 313–323.
- Bowman, S. A., Gortmaker, S. L., Ebbeling, C. B., Pereira, M. A., & Ludwig, D. S. (2004). Effects of fast-food consumption on energy intake and diet quality among children in a national household survey. *Pediatrics*, *113*(1), 112–118. Retrieved from <http://pediatrics.aappublications.org/content/113/1/112>
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101.
- Coryn, C. L., Noakes, L. A., Westine, C. D., & Schröter, D. C. (2011). A systematic review of theory-driven evaluation practice from 1990 to 2009. *American Journal of Evaluation*, *32*(2), 199–226.
- Deci, E. L., & Ryan, R. M. (2008). Self-determination theory: A macrotheory of human motivation, development, and health. *Canadian Psychology/Psychologie Canadienne*, *49*(3), 182–185.

- Donaldson, S. I., & Lipsey, M. W. (2006). Roles for theory in contemporary evaluation practice: Developing practical knowledge. In I. Shaw, J. C. Greene, & M. M. Mark (Eds.), *The handbook of evaluation: Policies, programs, and practices* (pp. 56–75). London, UK: Sage.
- Gordon, E. S., Tucker, P., Burke, S. M., & Carron, A. V. (2013). Effectiveness of physical activity interventions for preschoolers: A meta-analysis. *Research Quarterly for Exercise and Sport*, 84(3), 287–294.
- Greenhalgh, T., Wong, G., Westhorp, G., & Pawson, R. (2011). *Protocol – Realist and meta-analysis synthesis: Evolving Standards (RAMESES)*. Retrieved from <http://bmcmmedresmethodol.biomedcentral.com/articles/10.1186/1471-2288-11-115>
- Hills, A. P., King, N. A., & Armstrong, T. P. (2007). The contribution of physical activity and sedentary behaviors to the growth and development of children and adolescents. *Sports Medicine*, 37(6), 533–545. Retrieved from <https://link-springer-com.cyber.usask.ca/content/pdf/10.2165%2F00007256-200737060-00006.pdf>
- Kerlinger, F. N. (1986). *Foundations of behavioral research* (3rd ed.). New York, NY: Holt, Rinehart and Winston.
- Linsley, P., Howard, D., & Owen, S. (2015). The construction of context-mechanisms-outcomes in realistic evaluation. *Nurse Researcher* (2014+), 22(3), 28–38.
- Marchal, B., Van Belle, S., Van Olmen, J., Hoerée, T., & Kegels, G. (2012). Is realist evaluation keeping its promise? A review of published empirical studies in the field of health systems research. *Evaluation*, 18(2), 192–212.
- Maxwell, J. A. (2012). *A realist approach for qualitative research*. London: Sage.
- May, C., & Finch, T. (2009). Implementation, embedding, and integration: An outline of normalization process theory. *Sociology*, 43(3), 535–554.
- Merton, R. (1968). *Social theory and social structure*. New York, NY: Free Press.
- NVivo qualitative data analysis software. (2018). *QSR International Pty Ltd. Version 11*.
- Pawson, R. (2006). *Evidence based practice – A realist perspective*. London: Sage.
- Pawson, R. (2010). Middle range theory and program theory evaluation: From provenance to practice. In J. Vaessen & F. L. Leeuw (Eds.), *Mind the gap: Perspectives on policy evaluation and the social sciences* (pp. 171–202). New Brunswick, NJ: Transaction Publishers.
- Pawson, R., & Tilley, N. (1997). *Realistic evaluation*. London, UK: Sage.
- Pawson, R., & Tilley, N. (2004). *Realist evaluation*. Retrieved from http://www.communitymatters.com.au/RE_chapter.pdf
- Rogers, P. J. (2008). Using programme theory to evaluate complicated and complex aspects of interventions. *Evaluation*, 14(1), 29–48.
- Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55(1), 68–78.
- Salter, K. L., & Kothari, A. (2014). Using realist evaluation to open the black box of knowledge translation: A state-of-the-art review. *Implementation Science*, 9(1), 115.
- Sayer, A. (1992). *Method in social science – A realist approach*. London, UK: Routledge.
- Shearn, K., Allmark, P., Piercy, H., & Hirst, J. (2017). Building realist program theory for large complex and messy interventions. *International Journal of Qualitative Methods*, 16(1), 1–11.
- Statistics Canada. (2015). *Table 105-2026. Measured children and youth body mass index (BMI) (Cole classification), by age group and sex, Canada and provinces*. Canadian Community Health Survey - Nutrition. Ottawa, ON.
- Temple, M., & Robinson, J. C. (2014). A systematic review of interventions to promote physical activity in the preschool setting. *Journal for Specialists in Pediatric Nursing*, 19(4), 274–284.
- Van Belle, S. B., Marchal, B., Dubourg, D., & Kegels, G. (2010). How to develop a theory-driven evaluation design? Lessons learned from an adolescent sexual and reproductive health programme in West Africa. *BMC Public Health*, 10(1), 741–751.
- Weiss, C. (1997). *Theory-based evaluation: Past, present, and future. New directions for evaluation*. San Francisco, CA: Jossey-Bass.
- Wong, G., Westhorp, G., Greenhalgh, J., Manzano, A., Jagosh, J., & Greenhalgh, T. (2017). Quality and reporting standards, resources, training materials and information for realist evaluation: The RAMESES II project. *Health Services and Delivery Research*, 5(28), 1–108.