

Original Paper

Identification of Complex Health Interventions Suitable for Evaluation: Development and Validation of the 8-Step Scoping Framework

Rosemary Davidson¹, PhD; Gurch Randhawa¹, PhD; Stephanie Cash², MA

¹Institute for Health Research, University of Bedfordshire, Luton, United Kingdom

²Flying Start, Luton, United Kingdom

Corresponding Author:

Rosemary Davidson, PhD

Institute for Health Research

University of Bedfordshire

University Square

Luton, LU1 3JU

United Kingdom

Phone: 44 (0)1234 400 400

Email: rosemary.davidson@beds.ac.uk

Abstract

Background: There is extensive literature on the methodology of evaluation research and the development and evaluation of complex interventions but little guidance on the formative stages before evaluation and how to work with partner organizations that wish to have their provision evaluated. It is important to be able to identify suitable projects for evaluation from a range of provision and describe the steps required, often with academic institutions working in partnership with external organizations, in order to set up an evaluation. However, research evaluating programs or interventions rarely discusses these stages.

Objective: This study aimed to extend work on evaluability assessment and pre-evaluation planning by proposing an 8-Step Scoping Framework to enable the appraisal of multiple programs in order to identify interventions suitable for evaluation. We aimed to add to the literature on evaluability assessment and more recent evaluation guidance by describing the processes involved in working with partner organizations.

Methods: This paper documents the steps required to identify multiple complex interventions suitable for process and outcome evaluation. The steps were developed using an iterative approach by working alongside staff in a local government organization, to build an evidence base to demonstrate which interventions improve children's outcomes. The process of identifying suitable programs for evaluation, thereby establishing the pre-evaluation steps, was tested using all Flying Start provision.

Results: The 8-Step Scoping Framework was described using the example of the local government organization Flying Start to illustrate how each step contributes to finding projects suitable for process and outcome evaluation: (1) formulating overarching key questions that encompass all programs offered by an organization, (2) gaining an in-depth understanding of the work and provision of an organization and engaging staff, (3) completing a data template per project/program offered, (4) assessing the robustness/validity of data across all programs, (5) deciding on projects suitable for evaluation and those requiring additional data, (6) negotiating with chosen project leads, both within and outside the organization, (7) developing individual project evaluation protocols, and (8) applying for ethical approval from the university and partner organization.

Conclusions: This paper describes the processes involved in identifying suitable projects for evaluation. It adds to the existing literature on the assessment of specific programs suitable for evaluation and guidance for conducting evaluations by establishing the formative steps required to identify suitable programs from a range of provision. This scoping framework particularly relates to academic partners and organizations tasked with delivering evidence-based services designed to meet local needs. The steps identified have been described in the context of early years provision but can be applied to a range of community-based evaluations, or more generally, to cases where an academic partner is working with external stakeholders to identify projects suitable for academic evaluation.

(*JMIR Res Protoc* 2019;8(3):e10075) doi: [10.2196/10075](https://doi.org/10.2196/10075)

KEYWORDS

complex interventions; early years; evaluation; multistakeholder provision

Introduction

There is extensive literature on evaluation research methodology and development and evaluation of complex interventions, from identifying existing evidence to measuring outcomes and understanding processes [1-7]. However, there is little guidance on the formative stages of identifying suitable services/programs for evaluation and the ways to work with partner organizations that wish to have their provision evaluated in order to build an evidence base related to their particular local, geographical or cultural context beyond basic advice [8-10]. A possible disadvantage of conducting a robust evaluation is the risk of finding no change or negative results, potentially influencing future funding decisions and reputations.

Research on program evaluation rarely discusses the steps involved prior to evaluation in order to identify suitable projects, often with academic institutions working in partnership with external organizations to set up an evaluation. Guidance assumes that projects have already been identified, providing detailed instructions to plan and conduct evaluations. For example, Newcomer et al assumed that the projects to be evaluated were already chosen, and evaluators and organization staff had planned their evaluation approach [11]. They described the fundamental considerations that evaluators and organizations should address before beginning any evaluation activities, starting with matching evaluation approach to key questions, producing methodological rigor and appropriate evaluation design, or identifying ways to apply an evaluation framework in a particular disciplinary context [12].

Pre-evaluation activities are mostly discussed in the literature on evaluability assessment, which was first conceptualized in the late 70s [13] after the costly, large-scale evaluations of major social interventions in the United States in that period reported no benefit. Poor evaluation approaches and ultimately, disappointing results, were thought to be the result of inadequate program definition and lack of development and specification of causal links between intervention actions and expected results. In response, a “pre-assessment of evaluability” was developed to improve evaluation methodology, not by assessing whether a program can be evaluated, “but [by] whether the program is ready to be managed to achieve desired performance and outcomes, what changes are needed to allow results-oriented management, and whether evaluation is likely to contribute to improved program performance” [14]. Evaluability assessment has been revived in recent years, because the demand for evidence-based practice of has increased [15].

Evaluability assessment is a systematic method to plan robust evaluations as well as a “low-cost pre-evaluation activity to prepare better for conventional evaluations of programmes” [16] in order to make sound decisions on evaluation methodology before funds are committed. The approach is viewed as a way to balance the growing demand for evidence through evaluation when limited resources are available [15]. A recent rapid scoping review showed the range of interventions

that have been assessed by evaluability assessment methodology to determine their suitability for evaluation, such as the State Asthma Programme [17], the Healthy Community Challenge Fund [18], and National Driver Retraining Programme [19].

Evaluability assessment focuses on the feasibility of evaluating a specific intervention and usually involves the following key stages: structured engagement with stakeholders to understand the context of a particular intervention and ensure evaluation findings are meaningful, development of a theory of change to inform implementation and identify key outcomes, review of existing literature and data to establish quality of evidence already available, and recommendations for proposed evaluation designs. Evaluability assessment allows researchers to assess the suitability of a specific intervention for evaluation by working through the aforementioned four stages.

As it is designed to assess the suitability of a particular intervention for evaluation, it assumes that organizations that want to have their provision evaluated have the expertise to identify projects suitable for evaluation. The 8-Step Scoping Framework detailed here guides researchers and stakeholders through the stages prior to evaluability assessment, where the appetite for evaluation exists but the scope is ill defined. The 8 steps described are discussed using an example of a local government organization, Flying Start, to illustrate how each step contributes to the ultimate aim of identifying projects suitable for a process and outcome evaluation.

Flying Start [20] is part of Luton Borough Council, which is a part of the UK local government. It is a unitary local authority; as such, it provides all local services including health and social care, education, and learning. Flying Start aims to improve social, emotional, and health outcomes for children from the point of pregnancy to the age of 5 years. The importance of the early years and inequality in developmental outcomes is well documented [21-26]. Flying Start and the University of Bedfordshire are developing a process and outcome-evaluation framework to establish the efficacy of their multistakeholder provision, find evidence of what works, and ensure the provision offered meets local needs [27].

In this paper, the term “provision” most often refers to all the work of an organization to discern what it offers; “programs” or “services” are terms more likely used by an organization to describe the services they offer to the public or clients; and “intervention” is the more scientific term researchers favor to describe a project, program, or service that is subject to a process and outcome evaluation.

There is limited information in the literature about the steps required to identify suitable interventions before conducting an evaluation. This paper therefore aims (1) to extend work on evaluability assessment and pre-evaluation scoping by proposing an 8-Step Scoping Framework to be applied prior to evaluability assessment to enable the appraisal of multiple programs in order to identify interventions suitable for evaluation and (2) to add to the literature on evaluability assessment and more recent

evaluation guidance by describing the processes involved when working with partner organizations.

Methods

This paper documents the steps involved in identifying multiple complex interventions suitable for process and outcome evaluations. We developed an 8-Step Scoping Framework to identify complex health interventions. The framework guides the selection of suitable interventions for evaluation from a range of projects. Evaluability assessment allows in-depth appraisal of one project and is particularly important when considering the evaluation of larger, costly interventions before making a decision to commission an evaluation and begin detailed evaluation planning (Figure 1).

To refine these stages, an iterative approach was taken in order to develop the steps described below. The process was developed through regular meetings between an academic institution and stakeholders, which allowed a collaborative and reflexive process where researchers reported progress and were able to form a critical understanding of stakeholder priorities, and practical considerations were balanced with research objectives. A log was maintained to document progress during the process of identifying projects suitable for evaluation.

The 8-Step Scoping Framework was developed over an 11-month period through a series of meetings with Flying Start, Luton Borough Council, and associated stakeholders. Table 1 details the nature of the meetings, attendee numbers, affiliations, and their roles in the development of the pre-evaluation framework. The meetings were conducted concurrently during

the development of the framework. All meetings, except the Scoping Framework planning meetings, were ongoing as part of Flying Start's activities. The Flying Start staff meetings allowed researchers to gain an in-depth understanding of Flying Start provision, range, number, and development of services. The Partnership Board and Project Evaluation Group meetings allowed the input of a range of professionals on evaluation scoping strategy, identification of framework steps, and feedback on framework development through various iterations. The Scoping Framework-planning meetings were devoted to reporting framework progress and allowed researchers to apply the developing scoping criteria to Flying Start projects.

Formulation of the steps was led by the following key aims: to determine ways to obtain a full understanding of the provision offered by an organization; to arrive at a consensus on the type of questions to be asked in order to assess suitability of provision for evaluation; to find the best way to obtain such information from stakeholders; to refine the scoping process to allow a decision on suitable projects, space for negotiation with project leads, and development of stand-alone project-evaluation protocols per chosen project.

Meeting minutes with decisions made were typed up and circulated for comment and discussion as the framework steps were defined and clarified. The process of identifying suitable programs for evaluation, thereby establishing the 8 scoping steps, was tested using all Flying Start provision. During the framework-development period, 36 programs/services were offered by Flying Start to families in Luton. First and successive drafts of the Scoping Framework were presented to Flying Start, council staff, researchers, and associated stakeholders over time.

Figure 1. Context of 8-Step Scoping Framework in relation to evaluability assessment and evaluation planning.

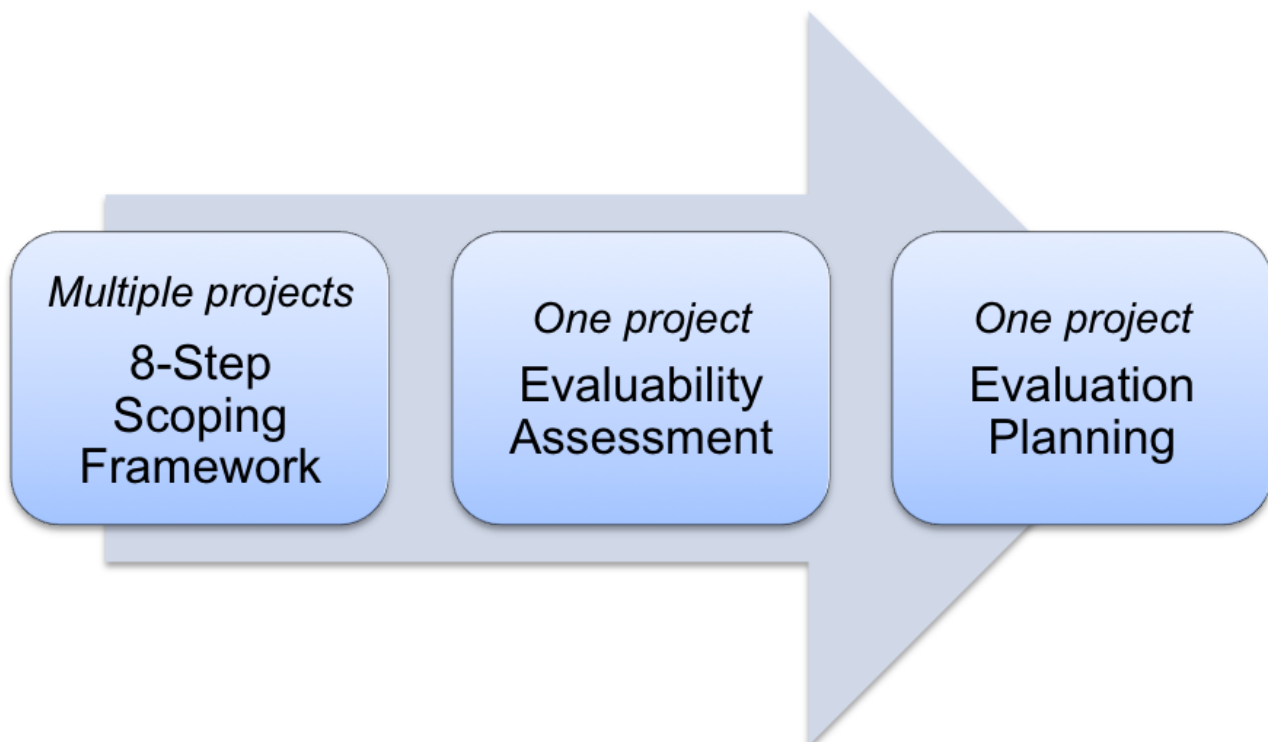


Table 1. Characteristics of the meetings used to develop the 8-Step Scoping Framework.

Characteristic	LBC ^a and Flying Start Partnership Board	Flying Start Project-Evaluation Group	Flying Start staff	Scoping Framework planning
Number of meetings held over the framework-development period	5 (bimonthly)	8 (every 4-5 weeks)	10 (monthly)	16 (every 2-3 weeks)
Purpose	Multiagency meeting to discuss issues related to early years services	Forum to discuss evaluation approaches for Flying Start services	Staff to update on progress and discuss any arising matters	Mapping of Scoping Framework progress
Attendees	Council heads of services, Flying Start staff, early years and public sector organizations, midwives, nutritionists, general practitioners, and councilors	Flying Start staff, LBC staff, and University of Bedfordshire staff	Flying Start staff, practitioners working in early years services, and university evaluation team	Flying Start senior staff and university evaluation team
Approximate number of attendees	15-20 stakeholders	8-12	15-20	3-5
Role in framework development	Input from a range of professionals and feedback on framework development	Identifying framework steps through various iterations	Gaining in-depth understanding of Flying Start provision, range, and number of services	Applying scoping criteria to all Flying Start services/programs offered

^aLBC: Luton Borough Council.

Results

The steps in the 8-Step Scoping Framework are presented in [Figure 2](#). The steps are discussed using a specific example of the local government organization Flying Start to illustrate how each step contributes to the ultimate aim of finding projects suitable for process and outcome evaluation.

Step 1 – Formulating Overarching Key Questions That Encompass All Programs Offered by an Organization

This may sound like an obvious first step, but it is important to determine whether the organization has an overarching aim guiding the content and purpose of their provision. Can this be translated into a research question to guide an evaluation? When considering multiple projects, is there a coherent research question that encompasses all projects? Most organizations will have key aims or a mission statement that can be reframed as a research question, which serves as a useful guide to ensure the overall evaluation strategy retains its focus and that the research question aligns with the objectives of the organization. In the case of Flying Start, a part of the Luton Borough Council, the overarching research question to guide the evaluation strategy was “What impact does an integrated early year’s strategy make on a life ready for learning at age five in a unitary authority?” This question was divided further into three subquestions:

- Has Flying Start succeeded in improving child/family outcomes?
- Was Flying Start more successful with certain groups and why?
- What aspects of Flying Start did participants (families) find most beneficial?

Irrespective of whether the organization has specific programs it wishes to evaluate or is led in consultation with an academic partner, the next step is vital.

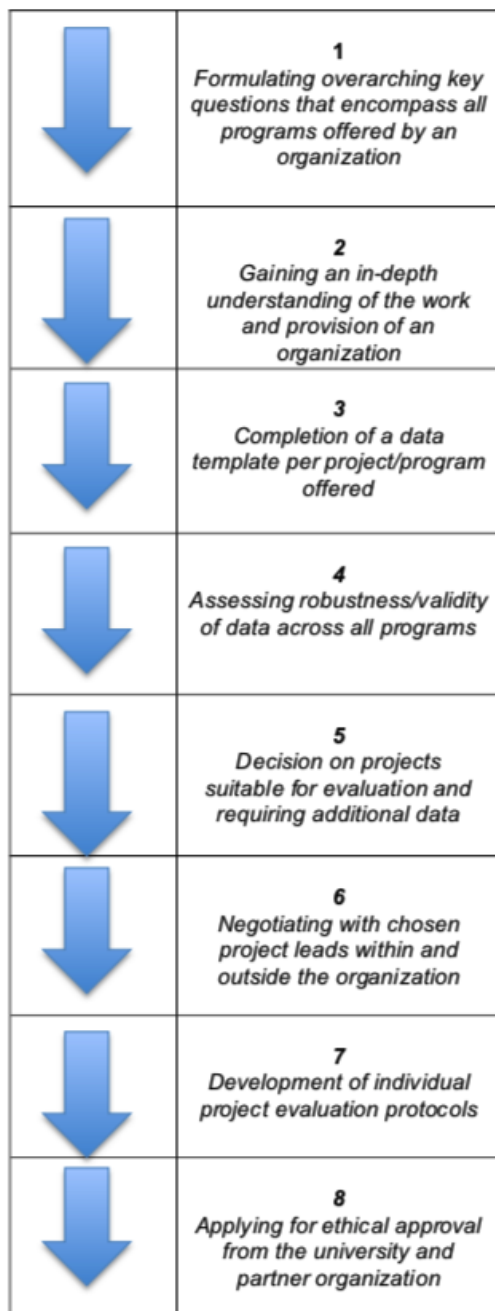
Step 2 – Gaining an In-Depth Understanding of the Work and Provision of an Organization

In the case of Flying Start, the requirement to evaluate their provision was a priority, but what was to be evaluated was unclear. Through a series of meetings, a logic model was developed to identify a set of questions that would help to both understand the provision and assess its suitability for evaluation [28]. What services would lend themselves to a robust process and outcome evaluation, which could then be published in peer-reviewed academic journals, thereby building a credible evidence base for their work? How to be strategic with the evaluation, given the finite resources and research capacity? In order to answer these questions, it is essential to negotiate access to the organization and be available to attend meetings, particularly where staff are given a forum to discuss their work and current progress. What may seem to be one organization from the outset is, in fact, a complex structure consisting of staff working in a considerable range of ways to deliver services. The researcher can begin understanding in detail how provision fits together, who it is aimed at, and the level of need it attempts to address.

Underlying all these issues, however, is an understanding of the pressures an organization faces, such as lack of staff, limited resources, responding to diverse and changing needs of a community, government guidelines, policy steers with ebbs and flows in funding, getting services up and running, and reaching those who need help most but are least likely to access services. Attending meetings and building relationships cement trust [1] and allays fears around being the subject of an evaluation. If a researcher is available to answer questions and give advice more generally about research and evaluation, it is possible to be a

valuable resource to the organization in terms of embedding interventions chosen for evaluation. evaluation methodology across all services, not only the

Figure 2. The 8-Step Scoping Framework for identifying complex health interventions suitable for evaluation.



Step 3 – Completion of a Data Template per Project/Program Offered

Understanding the range and content of an organization’s provision is different from establishing suitability for evaluation. With Flying Start, a template was developed for project/service leads to complete requesting information on target audience for service; service aims; whether any baseline data was collected prior to inception; key performance indicators linked to outcomes (eg, communication and language, nutrition and diet, and social and emotional development); the data available and what it signifies; who owns the data (is it in the public domain?);

who the data custodian is; how often data is collected; opportunities for tracking cohorts, size, and scale of the intervention; coverage (population, town wide, or ward level); whether participants (parents/families) are likely to be involved in more than one program; and the start date and length of intervention ([Multimedia Appendix 1](#)).

Researchers also agreed on key terminology during the process to avoid misunderstandings. For example, the term “provision” was used to describe the range of work an organization does to discern what was offered; “programs” or “services” were used to describe what was offered to the public or clients; and

“intervention” was the scientific term used to describe a project, program, or service that lends itself to a process and outcome evaluation.

An evaluation workshop was then organized with all project leads and other key Flying Start staff with the aim of encouraging staff to think about ways to evaluate their provision, answer evaluation and research questions, and provide guidance on how to fill in the data templates. A total of 36 programs/services were offered by Flying Start to families in Luton. The staff divided them into 8 domains: Antenatal/baby, Communication, Life course approach to healthy weight, Child mental health, Strong and supportive parents, System changes, Child safety, and Other. The event helped engage stakeholders in meaningful ways [29], foster partnerships, agree on the remit of different programs [6], support staff on ways to build an evidence base demonstrating the efficacy of their provision, and offer overall criteria to refer to in order to understand why some provision may be more suitable for evaluation. Possession of the completed data templates allowed different projects to be assessed and compared, leading to the next step.

Step 4 – Assessing Robustness/Validity of Data Across All Programs

With the completed data templates, the task of assessing suitability of projects for evaluation could be approached systematically. It was easier to approach project leads with any follow-up questions after referring to the templates. Key criteria (Table 2) for deciding on projects were assessment of quality of data (ie, use of validated outcome tools/scales for data collection); presence of an explicit statement on the causal assumptions of how the intervention will work [1]; presence of a robust existing theory underpinning the intervention [29]; whether a theory can be identified or developed if no theory is evident [2]; presence of any other factors that drive the program, such as experience and professional practice [1]; whether the intervention is developed to a point that it can reasonably be expected to have a worthwhile effect [2]; whether there are systems or protocols already built in to projects to collect the process and outcome data required; and whether it is possible to make comparisons with control groups in order to measure progress of those using a service or participating in an intervention. Practical considerations were paramount; for example, had a project already started [9]? How long would the

project/intervention run for? What were the funding restrictions? Strategically, a project may appear to be on a smaller scale, but may run for a sufficient amount of time to produce multiple cohorts of participants and therefore yield the quantity and depth of data required.

Services that appear suitable for evaluation on first inspection may, in fact, be in the early stages of implementation or facing implementation problems such as issues with recruitment or referral processes. Such problems are particularly significant when working with vulnerable families to, for example, assess the level of support required and willingness to engage with or attend services. Strategic decisions may have to be made to focus attention and resources elsewhere if an otherwise suitable program is facing problems with, for example, implementation, recruitment, or referrals of suitable participants. An exception may be made if a decision is taken to focus only on an early stage process study of an intervention that may not take off but may have strategic importance to an organization and contribute to academic debate. Such negotiations are most constructive when the preceding steps have been followed, allowing for face-to-face discussions and fruitful working relationships.

Step 5 – Decision on Projects Suitable for Evaluation and Requiring Additional Data

By applying the abovementioned criteria in step 4, the projects lending themselves to evaluation were identified. Research capacity [1] was then used as a guiding factor to ascertain what was possible to take on, by producing an evaluation timetable with timelines for each project under evaluation. As Flying Start offers a wide range of early years provision, it was also important, where practicably possible, to reflect diversity in the projects chosen. The provision/interventions chosen were Sign 4 Little Talkers/Big Feelings [30,31], which uses sign language to support the development of language, vocabulary, and positive behavior in children below 5 years of age; Healthy Exercise Nutrition for the Really Young (HENRY) [32-34], an obesity-prevention program for families with children below 5 years of age; Incredible Years [35-37], a parenting program for high risk socioeconomically disadvantaged families with children aged 3-5 years old having behavioral problems; and Parents as Partners [38,39], which offers counselling to improve couple relationships in order to improve child well-being and developmental outcomes (Table 3).

Table 2. Criteria for assessing data related to projects/interventions.

Criteria	Sign 4	HENRY ^a	Incredible Years	Parents as Partners
Data quality/outcome data	Sufficient	Sufficient	Sufficient	Sufficient
Theoretical basis	Yes	Yes	Yes	Yes
In-built evaluation tools	Yes	Yes	Yes	Yes
Prior evidence of positive effect	Pilot data	Yes	Yes	Yes
Control group comparisons	Retrospective and baseline data	Baseline data	Baseline data	Baseline data
Has the project started?	Yes	Yes	Yes	Under negotiation
Funding terms	Funded	Funded	Funded	Under negotiation
Is it scalable or does it involve multiple cohorts?	Yes	Yes	Yes	Yes

^aHENRY: Healthy Exercise Nutrition for the Really Young.

Table 3. Complex interventions identified by application of the 8-Step Scoping Framework.

Project characteristics	Sign 4	HENRY ^a	Incredible Years	Parents as Partners
Scope	<ul style="list-style-type: none"> To improve vocabulary and communication in preschool children 	<ul style="list-style-type: none"> Obesity prevention for parents of preschool children 	<ul style="list-style-type: none"> To address early onset behavioral problems in preschool children 	<ul style="list-style-type: none"> To improve couple relationship quality impacting children's outcomes
Aims	<ul style="list-style-type: none"> To investigate the impact of Sign 4 on early years outcomes and implementation, and lay and professional views 	<ul style="list-style-type: none"> To investigate pre- and postintervention impact on self-reported outcomes and implementation, and lay and professional views 	<ul style="list-style-type: none"> To investigate pre- and postintervention impact on self-reported outcomes and implementation, and lay and professional views 	<ul style="list-style-type: none"> To investigate pre- and postintervention impact on self-reported outcomes and implementation, and lay and professional views
Number of participants (n)	<ul style="list-style-type: none"> Preschool children (1500) Parents (20) Staff (30) Stakeholders (5) 	<ul style="list-style-type: none"> Parents (200) Facilitators (10) Stakeholders (5) 	<ul style="list-style-type: none"> Parents (140) Facilitators (10) Stakeholders (5) 	<ul style="list-style-type: none"> Parents (100) Facilitators (12) Stakeholders (5)
Data type	<ul style="list-style-type: none"> Early years outcomes, well-being scales, interviews 	<ul style="list-style-type: none"> Self-report measures and interviews 	<ul style="list-style-type: none"> Self-report measures, parenting questionnaires, and interviews 	<ul style="list-style-type: none"> Self-report measures, parenting questionnaires, and interviews

^aHENRY: Healthy Exercise Nutrition for the Really Young.

These projects already have systems in place to collect outcome data pre- and postintervention, with some opportunities to compare outcome data with existing larger datasets. Two of the four chosen are established programs running elsewhere in the United Kingdom or internationally, with published evidence demonstrating improved outcomes.

In terms of conducting evaluation research in Luton, we identified further cross-cutting questions resulting from the development of an in-depth knowledge of both provision and context, such as how established programs perform when implemented in highly ethnically and culturally diverse populations and the extent to which these complex interventions can be tailored to local circumstances or allow a degree of adaptation [2]; whether Flying Start is able to replicate the positive results reported from pilot studies or improve on outcomes published elsewhere; and collecting qualitative process data to understand *how* such improved outcomes were achieved (or not), which is an aspect of particular value to Flying Start, given that projects initially tend to be rolled out on a small scale.

Consequently, an overarching process-evaluation model was developed, which could be applied and tailored, where appropriate, to all the projects to be evaluated:

- Observations of staff and facilitator training sessions
- Observations of intervention sessions with facilitators working with families and children
- Individual interviews (or focus groups, where deemed appropriate) with staff once interventions are running as well as with project leads and Flying Start leads/commissioners
- Interviews with families after completing the sessions with follow-up interviews at 6 and 12 months

The five steps described then lead to the sixth step, developing individual evaluation protocols.

Step 6 – Negotiating With Chosen Project Leads Both Within and Outside the Organization

In order to develop a separate, specific evaluation protocol per project, it was necessary to liaise with project leads from within

and outside Flying Start. From an academic point of view, the protocols were intended to stand alone as a plan to conduct a robust process and outcome evaluation; however, it was vital to receive regular feedback from Flying Start staff on what was achievable. This would include negotiating access to observe particular staff training sessions, meetings, mentoring, and shadowing routine visits and key program sessions with families; quantifying as precisely as possible the level of involvement required from all stakeholders named in each evaluation to allay the anxieties of stressed staff with challenging workloads; and actively listening to personnel involved at all levels about their concerns and aspects/dimensions of the programs that they particularly wanted to know more about, given their expertise of the local context and population demographics. While working in detail to map out the evaluation stages required for each intervention, it was necessary to remain aware of the organization's provision as a whole, specifically, the potential themes underlying all projects delivered.

In the case of Flying Start provision, a training course attended by a large proportion of Luton's early years workforce—Five to Thrive [40,41]—was of particular interest; this course coaches staff to apply evidence-based neuroscientific approaches in their practice to support families to strengthen attachment bonds by responding, talking, playing, relaxing, and cuddling their children. Representing a cornerstone of the Flying Start strategy, all study protocols included the aim of investigating the impact of this training on staff as a part of each process evaluation.

Step 7 – Development of Individual Project Evaluation Protocols

Protocol drafts were revised on numerous occasions, as Flying Start staff commented and questioned the evaluation approach and content. For particular projects where Flying Start had subcontracted part of the delivery to a partner organization offering the intervention, the protocols were also sent out externally for feedback and clarification as well as to academic colleagues based in other universities with prior/continuing involvement with the development of the original intervention or evaluations thereof. Once protocol drafts were approved, topic guide questions were formulated for the proposed process evaluation for each project. These questions were tailored for interviews with different stakeholders—Flying Start leads/commissioners, project leads, frontline staff/session facilitators, and families. Again, the draft questions were circulated to all Flying Start staff (and key external program staff, where appropriate) involved in the delivery of each project in order to draw upon their expertise and ensure key topics were explored sufficiently and no aspects were overlooked. With agreement on the content of the topic guides, information sheets, and consent forms, it was possible to apply for ethical approval.

Step 8 – Applying for Ethical Approval From the University and Partner Organization

Although the National Health Service in the United Kingdom, for example, has systems in place as a result of systematic and ongoing evaluations of health interventions, a local authority working with an academic partner is less common. Ethical approval was sought from both the University of Bedfordshire and Luton Borough Council. Applications had to fulfil

requirements of both the Institute for Health Research and Luton Borough Council for participant informed consent, data protection, and data storage. We worked with the Council's Information Governance Team to ensure the requirements of their Tier 3 Information Sharing Agreement, detailing data-sharing processes for each Flying Start project under evaluation, were met as well as to write a master Information Sharing Agreement outlining the overarching principles all parties must adhere to as part of the evaluation research process. Additional time was required to ensure that the Council's data-sharing and informed consent guidance was met, which must also adhere to European Union law in this area. Finally, each research protocol was registered in the ISRCTN (International Standard Registered Clinical/social sTudy Number) registry to maximize awareness of the evaluations to other researchers, clinicians, and the public as well as to promote transparency and reduce duplication and selective reporting [27,42-44].

Discussion

Overview

This paper describes the preparation and work required to identify multiple complex projects/interventions suitable for process and outcome evaluation from a range of provision offered by an organization (in this case, services), designed to improve early years outcomes as part of local council provision in the United Kingdom. It details the complexities of academic partners working with a local authority to lay the foundation for a robust evaluation, with the aim of sharing this learning with others who are considering working within a similar model. We outline these steps in relation to previous guidance on conducting evaluations and the preassessment of specific interventions, namely, evaluability assessment, prior to embarking on evaluation research.

Our work adds to existing literature on evaluation methodology by setting out the steps required, particularly related to academic partners and organizations tasked with delivering services designed to meet local needs. After the 8 steps are completed, or the process is in the latter stages, a detailed evaluability assessment may be carried out. This may be particularly important if the projects identified are large scale, costly interventions requiring considerable resources to evaluate and pressure to produce conclusive results. Furthermore, an advantage of using the 8 steps prior to evaluability assessment is that many of the questions about an intervention's performance and expected outcomes have already been explored before a more detailed appraisal can be made about intervention management and performance. For smaller scale interventions, the use of the 8-Step Scoping Framework may be sufficient to allow progression to the evaluation-planning stage.

Limitations

The 8-Step Pre-evaluation Framework covers the early stages of evaluation planning to identify complex interventions suitable for evaluation. Therefore, this paper does not address economic aspects such as a cost-benefit analysis of late intervention [45,46] and how including such expertise may strengthen an evaluation and offer a business case for future funding or

commissioning decisions [47]. Further work should be undertaken to address how and when economic expertise would fit in to evaluation planning and how additional resources would be factored in to allow for this.

The 8 steps described are tested in relation to academic institutions working with local government in order to build an evidence base but is intended to be applied in other contexts where the goal is to develop a program of evaluation to identify what works. This could be, for example, healthcare, national government, or educational settings. Key criteria are that some form of program, provision, or service be offered with a defined purpose to change or improve a particular outcome(s). Therefore, at this stage, it could be argued that the 8-Step Scoping Framework may be applied in a wide variety of settings and contexts where academic evaluation is required. However, further refinement will likely be required, as others apply the 8-Step Framework and report on its generalizability and their experiences of identifying projects/interventions for evaluation.

The 8 steps are contingent upon an organization being open to having their provision evaluated and to change or modify procedures to ensure that data collection can take place. It requires researchers with good interpersonal and communication skills who are able to ask the pertinent questions and develop positive working relationships [47]. It is also important to note that the process of identifying suitable interventions is, in part, iterative and dependent on context, with some stages overlapping and feeding into each other in order to maintain momentum and ensure the most efficient use of time while considering a wide range of provision.

Evaluation Challenges, and Future Plans

The 8-Step Scoping Framework will be refined by continuing to work with Flying Start to identify further projects for evaluation in 2019 as well as seeking detailed feedback on how the organization has found the experience of working with an academic partner and being the subject of evaluation activities. As the evaluation of the chosen projects progresses, it will be possible to reflect further and refine the steps set out in this

paper. The ongoing impact of evaluation work is a dynamic process. As early results emerge, positive effects will reinforce original decisions to build an evidence base, whereas less conclusive or negative results may lead to skepticism and disappointment. Considering the weight of expectations around evaluation, a regular dialogue about the impact of results, coupled with a reminder of how process findings should help improve different aspects of provision, may help resolve any arising issues. We highlight the importance of developing positive working relationships and harnessing the expertise of organizations to ensure an evaluation asks the pertinent questions and explores the key issues.

Four Flying Start projects were found to be suitable for evaluation: Sign 4, HENRY, Incredible Years, and Parents as Partners. The selected projects fulfil our evaluation criteria to varying degrees: They collect key outcome data, allow comparisons with control groups, are established or imminent, are sizeable and scalable to allow for a mixed-methods approach, and use databases to allow tracking over time and have scope for inclusion of follow-ups. Our process and outcome-evaluation framework will enable us to assess what works and why it works. The steps identified have been described in the context of early years provision but can be applied to broader community-based evaluations. The process of formulating key evaluation questions in step 1 will ensure that the overarching evaluation strategy will retain its focus and we continue to be aligned with the objectives the organization.

Our subsequent evaluation of Flying Start provision must set realistic and achievable goals with the help of a detailed timetable, including contingency plans and a degree of slippage. Consideration must be given to issues of fidelity, whereby interventions may differ substantially between areas/settings, as projects need to adapt and take into account the needs of different communities. The evaluations must also consider difficult-to-reach groups who are less likely to access Flying Start services and that in the process of working with diverse communities, people do not fit into nicely packaged intervention and evaluation “boxes”.

Authors' Contributions

RD was the project coordinator for the study and collated completed data templates, assessed data quality, identified projects for evaluation, liaised with project leads, developed and drafted individual project-evaluation protocols, applied for ethical approval, and drafted the manuscript. GR was the principal investigator and provided oversight for the study and comments on the entire manuscript. SC provided oversight for the study, designed the data templates, organized the evaluation workshop, and provided comments and edits on the entire manuscript.

Conflicts of Interest

RD received funding from the University of Bedfordshire and Flying Start Luton, UK, for this study.

Multimedia Appendix 1

Data template for the project/program offered.

[\[PDF File \(Adobe PDF File\), 38KB-Multimedia Appendix 1\]](#)

References

1. Moore GF, Audrey S, Barker M, Bond L, Bonell C, Hardeman W, et al. Process evaluation of complex interventions: Medical Research Council guidance. *BMJ* 2015;350:h1258 [[FREE Full text](#)] [Medline: [25791983](#)]

2. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M. Developing and evaluating complex interventions: The new Medical Research Council guidance. *Int J Nurs Stud* 2013 May;50(5):587-592. [doi: [10.1016/j.ijnurstu.2012.09.010](https://doi.org/10.1016/j.ijnurstu.2012.09.010)] [Medline: [23159157](https://pubmed.ncbi.nlm.nih.gov/23159157/)]
3. Campbell NC, Murray E, Darbyshire J, Emery J, Farmer A, Griffiths F, et al. Designing and evaluating complex interventions to improve health care. *BMJ* 2007 Mar 03;334(7591):455-459 [FREE Full text] [doi: [10.1136/bmj.39108.379965.BE](https://doi.org/10.1136/bmj.39108.379965.BE)] [Medline: [17332585](https://pubmed.ncbi.nlm.nih.gov/17332585/)]
4. Thomson H, Hoskins R, Petticrew M, Ogilvie D, Craig N, Quinn T, et al. Evaluating the health effects of social interventions. *BMJ* 2004 Jan 31;328(7434):282-285 [FREE Full text] [doi: [10.1136/bmj.328.7434.282](https://doi.org/10.1136/bmj.328.7434.282)] [Medline: [14751903](https://pubmed.ncbi.nlm.nih.gov/14751903/)]
5. Campbell M, Fitzpatrick R, Haines A, Kinmonth AL, Sandercock P, Spiegelhalter D, et al. Framework for design and evaluation of complex interventions to improve health. *BMJ* 2000 Sep 16;321(7262):694-696 [FREE Full text] [Medline: [10987780](https://pubmed.ncbi.nlm.nih.gov/10987780/)]
6. Milstein B, Wetterhall S. A Framework Featuring Steps and Standards for Program Evaluation. *Health Promotion Practice* 2016 Jun 30;17(3):221-228 [FREE Full text] [doi: [10.1177/152483990000100304](https://doi.org/10.1177/152483990000100304)]
7. Bradley F, Wiles R, Kinmonth AL, Mant D, Gantley M. Development and evaluation of complex interventions in health services research: case study of the Southampton heart integrated care project (SHIP). The SHIP Collaborative Group. *BMJ* 1999 Mar 13;318(7185):711-715 [FREE Full text] [Medline: [10074018](https://pubmed.ncbi.nlm.nih.gov/10074018/)]
8. Posavac E. *Program Evaluation: Methods and Case Studies*. UK: Routledge; 2015.
9. Davidson JE. *Evaluation Methodology Basics: The Nuts and Bolts of Sound Evaluation*. Thousand Oaks: SAGE Publications; 2004.
10. Weiss C. *Evaluation: Methods for studying programs and policies*. Cape Town: Prentice Hall; 1997.
11. Newcomer KE, Hatry HP, Wholey J. Planning and Designing Useful Evaluations. In: *Handbook of Practical Program Evaluation*. New Jersey: Jossey-Bass; 2015:5-29.
12. Kaufman D, Roberts WD, Merrill J, Lai T, Bakken S. Applying an evaluation framework for health information system design, development, and implementation. *Nurs Res* 2006;55(2 Suppl):S37-S42. [Medline: [16601633](https://pubmed.ncbi.nlm.nih.gov/16601633/)]
13. Wholey JS. *Evaluation: Promise and Performance*. Washington DC: Urban Institute; 1979.
14. Strosberg MA, Wholey JS. Evaluability assessment: from theory to practice in the Department of Health and Human Services. *Public Adm Rev* 1983;43(1):66-71. [Medline: [10259244](https://pubmed.ncbi.nlm.nih.gov/10259244/)]
15. Craig P, Campbell M. Evaluability Assessment: a systematic approach to deciding whether and how to evaluate programmes and policies. Glasgow: What Works Scotland; 2015. URL:<http://whatworksscotland.ac.uk/wp-content/uploads/2015/07/WWS-Evaluability-Assessment-Working-paper-final-June-2015.pdf> [accessed 2018-12-10] [WebCite Cache ID 74Yn7eNeE]
16. Leviton LC, Khan LK, Rog D, Dawkins N, Cotton D. Evaluability assessment to improve public health policies, programs, and practices. *Annu Rev Public Health* 2010;31:213-233. [doi: [10.1146/annurev.publhealth.012809.103625](https://doi.org/10.1146/annurev.publhealth.012809.103625)] [Medline: [20235852](https://pubmed.ncbi.nlm.nih.gov/20235852/)]
17. Hester LL, Wilce MA, Gill SA, Disler SL, Collins P, Crawford G. Roles of the state asthma program in implementing multicomponent, school-based asthma interventions. *J Sch Health* 2013 Dec;83(12):833-841 [FREE Full text] [doi: [10.1111/josh.12101](https://doi.org/10.1111/josh.12101)] [Medline: [24261517](https://pubmed.ncbi.nlm.nih.gov/24261517/)]
18. Ogilvie D, Cummins S, Petticrew M, White M, Jones A, Wheeler K. Assessing the evaluability of complex public health interventions: five questions for researchers, funders, and policymakers. *Milbank Q* 2011 Jun;89(2):206-225 [FREE Full text] [doi: [10.1111/j.1468-0009.2011.00626.x](https://doi.org/10.1111/j.1468-0009.2011.00626.x)] [Medline: [21676021](https://pubmed.ncbi.nlm.nih.gov/21676021/)]
19. Joannis M, Stinchcombe A, Yamin S. Evaluability assessment of a national driver retraining program: Are we evaluating in the right lane? *Can J Prog Eval* 2011;25(1):27-50.
20. Flying Start. URL:<https://www.flyingstartluton.com> [accessed 2018-05-11] [WebCite Cache ID 6zLGYoX2A]
21. Chowdry H, McBride T. Early Intervention Foundation. London; 2017. Disadvantage, behavior and cognitive outcomes: Longitudinal analysis from age 5 to 16 URL:<https://www.eif.org.uk/report/disadvantage-behaviour-and-cognitive-outcomes/> [accessed 2018-12-10] [WebCite Cache ID 74Yqcl1BP]
22. Allen G. Stationery Office: HM Government; 2011. Early intervention: the next steps URL:https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/284086/early-intervention-next-steps2.pdf [accessed 2018-12-10] [WebCite Cache ID 74YrMx9pt]
23. Ermisch J. Origins of Social Immobility and Inequality: Parenting and Early Child Development. *National Institute Economic Review* 2008 Jul;205(1):62-71. [doi: [10.1177/0027950108096589](https://doi.org/10.1177/0027950108096589)]
24. Melhuish E, Belsky J, Leyland AH, Barnes J, National Evaluation of Sure Start Research Team. Effects of fully-established Sure Start Local Programmes on 3-year-old children and their families living in England: a quasi-experimental observational study. *Lancet* 2008 Nov 08;372(9650):1641-1647 [FREE Full text] [doi: [10.1016/S0140-6736\(08\)61687-6](https://doi.org/10.1016/S0140-6736(08)61687-6)] [Medline: [18994661](https://pubmed.ncbi.nlm.nih.gov/18994661/)]
25. Blanden J. University College London. London: Corporate Document Services, Leeds; 2006. Bucking the Trend? What enables those who are disadvantaged in childhood to succeed later in life? URL:<http://dera.ioe.ac.uk/7729/1/WP31.pdf> [accessed 2018-12-10] [WebCite Cache ID 74YxSjFTA]
26. Feinstein L. Inequality in the Early Cognitive Development of British Children in the 1970 Cohort. *Economica* 2003 Feb;70(277):73-97. [doi: [10.1111/1468-0335.t01-1-00272](https://doi.org/10.1111/1468-0335.t01-1-00272)]

27. Davidson R, Randhawa G, Cash S. ISRCTN. 2018. Steps to Evaluation: How to set up an evaluation with external organisations URL:<http://www.isrctn.com/ISRCTN26468622> [accessed 2018-12-11] [WebCite Cache ID 74b8xAiX8]
28. McLaughlin JA, Jordan GB. Logic models: a tool for telling your programs performance story. *Evaluation and Program Planning* 1999 Mar;22(1):65-72. [doi: [10.1016/S0149-7189\(98\)00042-1](https://doi.org/10.1016/S0149-7189(98)00042-1)]
29. Goodman RM. Principals and tools for evaluating community-based prevention and health promotion programs. *J Public Health Manag Pract* 1998 Mar;4(2):37-47.
30. Sign4learning. URL:<http://www.sign4learning.co.uk> [accessed 2018-05-11] [WebCite Cache ID 6zLXWzdSI]
31. Cook SW, Mitchell Z, Goldin-Meadow S. Gesturing makes learning last. *Cognition* 2008 Feb;106(2):1047-1058 [FREE Full text] [doi: [10.1016/j.cognition.2007.04.010](https://doi.org/10.1016/j.cognition.2007.04.010)] [Medline: [17560971](https://pubmed.ncbi.nlm.nih.gov/17560971/)]
32. HENRY. URL:<https://www.henry.org.uk> [accessed 2018-05-12] [WebCite Cache ID 6zLYgE7jp]
33. Willis TA, Roberts KPJ, Berry TM, Bryant M, Rudolf MCJ. The impact of HENRY on parenting and family lifestyle: A national service evaluation of a preschool obesity prevention programme. *Public Health* 2016 Jul;136:101-108. [doi: [10.1016/j.puhe.2016.04.006](https://doi.org/10.1016/j.puhe.2016.04.006)] [Medline: [27184820](https://pubmed.ncbi.nlm.nih.gov/27184820/)]
34. Willis TA, George J, Hunt C, Roberts KPJ, Evans CEL, Brown RE, et al. Combating child obesity: impact of HENRY on parenting and family lifestyle. *Pediatr Obes* 2014 Oct;9(5):339-350. [doi: [10.1111/j.2047-6310.2013.00183.x](https://doi.org/10.1111/j.2047-6310.2013.00183.x)] [Medline: [23818487](https://pubmed.ncbi.nlm.nih.gov/23818487/)]
35. The Incredible Years. URL:<http://www.incredibleyears.com> [accessed 2018-05-12] [WebCite Cache ID 6zLaPO8Tb]
36. Barlow J, Smailagic N, Huband N, Roloff V, Bennett C. Group-based parent training programmes for improving parental psychosocial health. *Cochrane Database Syst Rev* 2012 Jun 13(6). [doi: [10.1002/14651858.CD002020.pub3](https://doi.org/10.1002/14651858.CD002020.pub3)] [Medline: [22696327](https://pubmed.ncbi.nlm.nih.gov/22696327/)]
37. Presnall N, Webster-Stratton CH, Constantino JN. Parent training: equivalent improvement in externalizing behavior for children with and without familial risk. *J Am Acad Child Adolesc Psychiatry* 2014 Aug;53(8):879-887 [FREE Full text] [doi: [10.1016/j.jaac.2014.04.024](https://doi.org/10.1016/j.jaac.2014.04.024)] [Medline: [25062595](https://pubmed.ncbi.nlm.nih.gov/25062595/)]
38. Tavistock Relationships. Parents as Partners URL:<https://tavistockrelationships.org/relationship-help/parents-as-partners-programme> [accessed 2018-05-24] [WebCite Cache ID 6zefyIE8F]
39. Harold G, Acquah D, Sellers R, Chowdry H, Feinstein L. Early Intervention Foundation. London; 2016. What works to enhance inter-parental relationships and improve outcomes for children URL:<https://www.eif.org.uk/report/what-works-to-enhance-interparental-relationships-and-improve-outcomes-for-children/> [accessed 2018-12-10] [WebCite Cache ID 74Z3Ty0pf]
40. Five to thrive. URL:<http://www.fivetothrive.org.uk> [accessed 2018-12-10] [WebCite Cache ID 74Z3hQtFK]
41. Cozolino L. *The Neuroscience of Human Relationships: Attachment and the Developing Social Brain*. New York: W.W. Norton & Company; 2014.
42. Davidson R, Randhawa G. Improving speech, language and communication in pre-school children: Evaluation of the Sign 4 Programme. 2018. URL:<http://www.isrctn.com/ISRCTN42025531> [accessed 2018-12-12] [WebCite Cache ID 74bAd73fA]
43. Davidson R, Randhawa G. Reducing obesity in pre-school children: Implementation and Effectiveness of the HENRY Programme, Luton, UK. 2018. URL:<http://www.isrctn.com/ISRCTN62454430> [accessed 2018-12-11] [WebCite Cache ID 74bAyH9Ms]
44. Davidson R, Randhawa G. Reduction and prevention of aggression and behavioural problems in disadvantaged pre-school children: Implementation and effectiveness of the Incredible Years Programme, Luton, UK. 2018. URL:<http://www.isrctn.com/ISRCTN14311503> [accessed 2018-12-12] [WebCite Cache ID 74bBOz9Jp]
45. Chowdry H, Fitzsimons P. Early Intervention Foundation. London; 2016. The cost of late intervention: EIF analysis 2016 URL:<https://www.eif.org.uk/report/the-cost-of-late-intervention-eif-analysis-2016> [accessed 2018-12-10] [WebCite Cache ID 74Z5LUA1I]
46. Chowdry H. Early Intervention Foundation. London: Early Intervention Foundation; 2015. Making an early intervention business case: Checklist and recommendations for cost-benefit analysis URL:<https://www.eif.org.uk/resource/making-an-early-intervention-business-case-checklist-and-recommendations-for-cost-benefit-analysis> [accessed 2018-12-10] [WebCite Cache ID 74Z6lPZZI]
47. Owen JM. *Program Evaluation: Forms and Approaches*. New York: The Guilford Press; 2006.

Abbreviations

HENRY: Healthy Exercise Nutrition for the Really Young

LBC: Luton Borough Council

Edited by G Eysenbach; submitted 08.02.18; peer-reviewed by B Chaudhry, E Ammenwerth, L Becker, H Durrani; comments to author 31.03.18; revised version received 24.05.18; accepted 10.11.18; published 05.03.19

Please cite as:

Davidson R, Randhawa G, Cash S

Identification of Complex Health Interventions Suitable for Evaluation: Development and Validation of the 8-Step Scoping Framework
JMIR Res Protoc 2019;8(3):e10075

URL: <https://www.researchprotocols.org/2019/3/e10075/>

doi: [10.2196/10075](https://doi.org/10.2196/10075)

PMID: [30835240](https://pubmed.ncbi.nlm.nih.gov/30835240/)

©Rosemary Davidson, Gurch Randhawa, Stephanie Cash. Originally published in JMIR Research Protocols (<http://www.researchprotocols.org>), 05.03.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on <http://www.researchprotocols.org>, as well as this copyright and license information must be included.