2020

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 774210

CO-CREATE

D7.4: A protocol for developing implementation and evaluation plans

University of Oslo
30.04.2020

Deliverable administration and summary			
Due date	30.04.2020		
Submission date	29.04.2020		
Deliverable type	Report		
Contributors:	**Name**	**Organisation**	**Role / Title**
Deliverable Leader	Nanna Lien	UoO	WP7-leader/Professor
Contributing Author(s)			
Reviewer(s)	Arnfinn Helleve Deanna Hoelscher Julianne Williams	NIPH UTHealth WHO Europe	
Final review and approval	Nanna Lien	UoO	WP7-leader/Professor

Document change history				
Version	**Release date**	**Reason for Change**	**Status (Draft/In-review/Submitted)**	**Distribution**

Dissemination level		
PU	Public	X

Executive Summary

This protocol is part of the Objective 7.2: *To develop implementation and evaluation plans for 1-3 selected co-created obesity-related policy interventions (tools, strategies, programmes) in each of the five countries* in Work package 7 in the CO-CREATE project, led by the University of Oslo.

Planning for implementation and evaluation of policies should be standard practice in public health in order to provide recommendations for termination, maintenance, improvements and or/scaling up of polices in a timely manner. However, lack of knowledge in and skills on how to develop and follow up implementation and evaluation plans are potential barriers to establishing such a practice. Building capacity for this through partnership between the CO-CREATE-partners and local stakeholders of a policy idea is one contribution to changing this. The protocol thus consists of two parts – firstly, an overview of evaluation and implementation of policy and secondly, the step by step protocol on how to develop implementation and evaluation plans for policy ideas from youth in the CO-CREATE project.

The 1-3 policy ideas that will be used to write the implementation and evaluation plans are developed by youth in the alliances as part of WP5 and discussed with relevant stakeholders as part of the dialogue forums in WP6. Within the five years of the CO-CREATE project, the enactment of the policy can not be promised. Yet, the process of writing these plans will contribute to maintaining interest in the idea, as well as building capacity for implementation and evaluation locally. This is likely to contribute to the likelihood of the policy ideas being enacted at some time. The enactment process will probably change the policy idea and thus the plans will be drafts that needs to be finalized when the policy has been enacted. Local policy ideas - i.e. in the community or specific settings - might be more feasible to write plans for than national policy ideas, and the guide is written for that purpose.

The protocol describes how to select the 1-3 policy ideas, establish a core team, draw logic models for implementation and evaluation, write up the implementation plan and focus the evaluation before summarizing it all in one implementation and evaluation plan. It is supported by templates for an evaluation framework, worksheets to be used during the process and for writing up the plans. Given that the types of policy ideas, the local context supporting the idea and the background knowledge in evaluation, implementation and system science varies greatly across the CO-CREATE partners, the protocol will have to be executed with flexibility and the support of the lead partner as well as the other partners in a collaborative learning process. This learning process can be used to improve the protocol.

Table of content

List of acronyms / abbreviations

CDC - Centers for Disease Control and Prevention

Center TRT - Center of Excellence for Training and Research Translation

IOM – Institute of Medicine

RCT - Randomized controlled trial

WHO – World Health Organization

Introduction

Work package 7 (WP7) has the overarching aim to evaluate the project using process, output and impact data. This aim is broken down into three objectives with corresponding tasks. This deliverable is part of Objective 7.2: *To develop implementation and evaluation plans for 1-3 selected co-created obesity-related policy interventions (tools, strategies, programmes) in each of the five countries*, and the corresponding Task 7.3 *To develop implementation and evaluation plans for 1-3 selected co-created obesity-related policy interventions (tools, strategies, programmes) within the five countries*. This task has the following three deliverables, and involves five of the CO-CREATE-partners (UoO, UvA, LSHTM, CEIDSS, SWPS):

D7.4: A protocol for developing implementation and evaluation plans (Month 24)
D7.5: A workshop for CO-CREATE co-workers on how to apply the protocol (Month 24)
D7.6: Implementation and evaluation plans (in local languages) from each of the five case countries (Month 54)

Planning for implementation and evaluation of policies should be standard practice in public health in order to provide recommendations for termination, maintenance, improvements and or/scaling up of polices in a timely manner. However, lack of knowledge in and skills on how to develop and follow up implementation and evaluation plans are potential barriers to establishing such a practice. Thus, building capacity for this through partnership between the CO-CREATE-partners and local stakeholders of a policy idea is one contribution to changing this. This deliverable is written to guide this process and consist of two parts – firstly, an overview of evaluation and implementation of policy and secondly, the step by step protocol on how to develop implementation and evaluation plans in the CO-CREATE project.

Deliverable description

In accordance with the grant agreement:

Evaluation plans will be based on obesity-related policy evaluation frameworks developed by World Health Organization (WHO) and by Institute of Medicine (IOM) in the US (WHO 2008, WHO 2013, IOM 2013) aimed at collecting process, output and impact data. Implementation plans should be based on relevant theories and frameworks (i.e. Nilsen 2015, Horodyska et al 2015) applied to the local context in systematic manner (Eldredge et al 2016). A protocol with an overall evaluation framework and a guide on developing implementation and evaluation plans will be developed. The CO-CREATE partners of the five countries will be trained and guided in the use of these protocols before developing implementation and evaluation plans for 1-3 selected co-created obesity-related policy interventions in their countries.

Relationship to WP5 and WP6

The policy ideas are developed by youth in the alliances (WP5) and discussed with the stakeholders (WP6). The ideas to be taken forward into implementation and evaluation plans should be those ideas that seem to be taken up by the stakeholders and are thus more likely to be enacted. The enactment of the policy is however not a requirement since that is dependent on political processes outside of CO-CREATE.

Relationship to D7.5 (workshop) and D7.6 (implementation and evaluation plans)

This protocol builds on the feedback received from the workshop (D7.5) where CO-CREATE partners were introduced to evaluation and implementation and tried out some of the proposed tools. The workshop and the protocol together with continued support from WP7 will enable the CO-CREATE partners to work with local stakeholders to develop evaluation and implementation plans. Based on the workshop, the protocol and the co-creation process, 1-3 plans per country will be delivered (D7.6). These plans are hypothetical and contingent on enactment of the policy ideas where upon the policy ideas might be changed and require updating and finalization of the plans before using them. The actual use of the plans will also depend on securing collaboration and additional funding beyond the task of delivering 1-3 implementation and evaluation plans per country.

Background

In order to tackle the obesity epidemic, there have been a call to move beyond interventions targeting the individual level to apply comprehensive packages of policies which address the epidemic as a result of an obesity system (Roberto et al 2015, IOM 2012, WHO 2008). However, there is a well-known challenge of lack of evaluation and incomplete implementation of policies, making it difficult to judge the contribution of comprehensive policies to solving the problem. Parts of the problem is that the gold standard for effect evaluation - the randomized controlled trial (RCT) - is not readily applicable for policy evaluations (IOM, 2013), and that process evaluation of the implementation processes in RCTs have usually been secondary to the main objective of effectiveness of the interventions.

According to Centers for Disease Control and Prevention (CDC), the policy process consists of five domains (problem identification, policy analysis, strategy and policy development, policy enactment and policy implementation) and the two continuous processes of stakeholder engagement and evaluation (CDC, 2012). Furthermore, they define policy evaluation as "*the systematic collection and analysis of information to make judgement about the context, activities, characteristics or outcomes of one or more domains of the policy process."* That is "*evaluation may inform and improve policy development, adoption, implementation and effectiveness and thus build the evidence for policy interventions*". Within public health, policy development "*includes advancing and implementing public health law, regulations, or voluntary practices that influence systems development,*

organizational change and individual behavior to promote improvements in health". An enacted policy could thus lead to one or several programs that aim to ensure that the objectives of the policy is achieved either at the national or community level. These programs would then be the objects of evaluation with regards to implementation and effectiveness.

WHO has provided a framework to monitor and evaluate implementation of their global strategy on diet, physical activity and health (WHO, 2008). However, the framework focuses on the national-level, which might not be a relevant or feasible first step for the policy ideas from the youth alliances in CO-CREATE. In the US, IOM has produced two reports - one describing how to accelerate obesity prevention (IOM 2012) and the other on how to evaluate the suggested policies and interventions both at the national and community level (IOM 2013). The latter report suggests how to set up a framework for evaluation - building on assessment by data to describe the problem, surveillance over time to track changes, monitoring of implementation of interventions and summative evaluation to attribute those changes to the interventions. The national level evaluation plan, is focusing on building leadership and infrastructure for surveillance and agreement on commonly used measurements of indicators. On the other hand, the community level evaluation plan consists of two parts - one for assessment and surveillance and one for monitoring and summative evaluation. Clearly, the national level and the building of infrastructure for surveillance at either level are important parts of an evaluation framework, but for the CO-CREATE project the primary focus will be on the monitoring and summative evaluation at the community level.

In this task of the CO-CREATE project, the focus will be on the continues process of **evaluation** of policy, as well as **implementation** of policies. Thus, the State-of-the art section will give an introduction to these two topics.

State-of-the art in the topic

1) Evaluation

Types of evaluation

In public health, three types of program evaluation are recognized; formative, process and outcome/impact evaluation (Bauman and Nutbeam, 2014). The same phases are recognized in policy evaluation and correspond to content, implementation and impact evaluation of the policy at the different phases of the CDC policy cycle.

Formative evaluation is the phase that *"determines the main problem and identify solutions that are feasible, appropriate, and meaningful for the target population"* (CDC, 2014). In public health program development, the output of the formative evaluation phase would be the program ready to

implement, since the program at that point has incorporated input through pilot or feasibility data obtained from relevant stakeholders and/or the priority population. The program should have been developed through applying theory- and evidence-based knowledge in combination with a thorough understanding of the determinants of the health problem in the target group as well as the opportunities for interventions within the local context (Eldredge et al, 2016). Such programs are usually multicomponent addressing both individual level beliefs, knowledge and skills (i.e. through a classroom component), as well as changing the environmental conditions through components targeting the change agents who have the power to change environmental factors determining the behaviors and/or health problems of the target group (i.e. parent involvement, a healthy canteen or active recess component). The possibility to pilot test the whole program as it is meant to be implemented has become more difficult with the increasing use of multi-component programs that should be implemented over longer periods/permanently. Nonetheless, pretesting critical components of the program for acceptability and feasibility by the target group and implementers is still of major importance in the formative evaluation phase (Eldredge et al 2016).

Process evaluation *"examines the implementation of the policy related activities"* (i.e. programs) (CDC, 2014). The process evaluation would be concerned with whether the activities were implemented as planned and the immediate outputs achieved. Process evaluations are also concerned with studying the effect of factors that are either barriers or facilitators of the implementation (Durlak & Dupre, 2008; Aarons et al 2011) to understand their impact on implementation outcomes (Proctor et al, 2011) and the subsequent effects of this on the effects studied in the outcomes/impact evaluation. A process evaluation could be formative or summative (Saunders at al, 2005). A ***formative process evaluation*** would feed the data back into the implementation process immediately in order to improve the implementation process. A ***summative process evaluation*** could collect data during as well as after the implementation process, but would not analyze the data until afterwards and the results would be primarily be used to understand the effects studied in the outcomes/impact evaluation, and based on this give recommendations for the future implementation of the policy activities.

The **outcome/impact evaluation** is concerned with determining if the policy (program/activities) achieved its objectives in the **short, intermediate, and long term (**Bauman and Nutbeam 2015, Figure 1.2, Leeman et al 2012). Short term impacts could be attitudes, knowledge or skills, as well as changes in organizational practices. Intermediate would typically be the behaviors of the target group and long term would be the health and societal effects. This type of evaluation is also called effectiveness evaluation. The gold standard for assessing effectiveness of interventions through research is the RCT. However, for evaluation of policies aimed at changing environmental conditions this might not be the best suited to answer the questions posed by the stakeholders and thus natural experiments or combining evidence from different sources in a summative evaluation might be better ways to assess outcome and impact (IOM, 2013).

Participatory evaluation is defined as *"a partnership approach to evaluation in which stakeholders actively engage in developing the evaluation and all phases of its implementation"* (Zukoski & Luluquisen, 2002). This is thus a different process from traditional research initiated evaluation

studies, and the methods used tend to be more qualitative although both qualitative and quantitative can be used (WHO 2013, Annex 7). The benefits of participatory evaluation are many:

- Identifying locally relevant evaluation questions
- Improve policy/program performance
- Empower participants
- Build capacity
- Develop leaders and build team
- Sustain organizational learning and growth

However, there are also the challenges of requiring *time and commitment* from many different stakeholders, securing *resources* and managing potential *conflicts* (Zukoski & Luluquisen, 2002). The evaluation might be less objective and the credibility might be questioned (WHO 2013, Annex 7).

Standards and steps to guide program evaluation

The CDC has developed a thorough framework to guide program evaluation in public health (CDC, 2014). The framework consists of four standards and six steps. The four standards to be considered throughout the process are: *utility* (who wants the results for what?), *feasibility* (of the procedures given the time, resources and expertise available), *propriety* (conducted fair and ethical?) and *accuracy* (given stakeholder needs and evaluation purpose). The six steps forming a cyclic process are:

1) Engaging stakeholders
2) Describe the program
3) Focus the evaluation design
4) Gather credible evidence
5) Justify conclusions
6) Ensure use and share lessons learned.

The three first steps will form the core steps for this protocol and are thus described in the following. However, step 6 should always be kept in mind as the aim of the evaluation should be to provide recommendations about the further use of the policy to the stakeholders (WHO, 2013).

1) Engaging stakeholders

The first step of Engaging the stakeholders is important in order to understand what kind of results they want from an evaluation and how they would like to use it. Stakeholders should at least include those funding the program, those running the program, the implementers and the target group, but there could also be other less obvious stakeholders. It is also important to listen to those that are opposed to the program.

Engaging the stakeholders is similarly stressed in making community obesity evaluation plans in the IOM report (IOM 2013, p 230) and when applying a systems perspective to evaluation (Hargreaves, 2010). The IOM stresses that community engagement provide changes through both specific programs, as well as policy, environment and system changes, and that there are feedback loops from both these changes which might cause increased community engagement over time (IOM 2013, Figure 8-1, page 232).

Stakeholders may be engaged in many different ways from one off conversations to including them in the core evaluation team or an advisory committee. The core evaluation team should consist of those who will be involved in the evaluation from planning to reporting, and should include the necessary content and evaluation methodology expertise in addition to key stakeholders. Access to expertise in practical policy implementation is also of particular importance. The team should be led by the lead evaluator who is overseeing the whole evaluation process.

2) Describing the program through use of logic models

The second step of Describing the program recommends using a logic model with Inputs, Activities, Outputs and Outcomes to visualize the program.

- ***Inputs*** are the information or resources required for developing and implementing the policy.
- ***Activities*** are the actions that are carried out in order to implement the policy.
- ***Outputs*** are the direct results of these action steps.
- ***Outcomes*** are the short-, intermediate and long-term changes in the environment, behaviors, social norms and health outcomes of the target group.

The CDC funded Center of Excellence for Training and research Translation (Center TRT) has developed an evaluation framework for obesity prevention policy interventions based on the CDC evaluation framework and with a logic model template as a core tool (Leeman et al 2012). The framework also draws on other theories and frameworks related to policy making and evaluation in public health, such as Kingdon's multiple streams theory of policy making and the public health impact criteria from the RE-AIM framework (Reach, Effectiveness, Adoption, Implementation and Maintenance) (Glasgow et al 2019). The multiple streams theory postulates that there are three separate streams; a problem, a policy and a politics stream and when these streams come together the policy is more likely to be enacted and implemented. The RE-AIM framework stresses the importance that all settings/sectors that ought to adopt an enacted policy should do so and ensure that it is implemented according to plan in order to reach the majority of the target population and thus the policy could exert its effectiveness in changing the population health. In addition, the maintenance of both program implementation and effects on the target population are important to monitor in order to evaluate the long-term impact of the policy. The website of Center TRT provides many resources including the logic model framework template and several complete examples of programs with their logic models and evaluation plans (http://centertrt.org/).

IOM propose a generic logic model for community-level initiatives to prevent obesity, where they differentiate between the program components that cause changes in the individual level believes and behaviors and the multi-faceted initiatives that changes policies, environments and systems (IOM 2013, Table 8-1, page 233). This might be a useful distinction when sorting out how the policy idea developed by the Youth Alliance will exert its effect in the local context.

Clearly, both Center TRT and IOM depict their logic models in very traditional, linear ways, but the IOM report does recommend taking a systems perspective also on evaluation (IOM 2013, Chapter 9), and one way to do this would be to start by drawing the logic model of the policy and it's implementation as systems models (Hargreaves 2010).

3) Focusing the evaluation design

The third step is Focus the evaluation design where agreements need to be reached on which evaluation questions are important and possible to answer within the available resources (time, money, expertise) and to write this up in an evaluation plan with clear assignment of roles and responsibilities of all involved.

- Who are the users and how will they use this evaluation?
- What kind of questions do they want answered – are they primarily on implementation or outcomes or both?
- What are the resources available?
- What are likely changes to be achieved within the timeframe envisaged for the evaluation?

The choice of designs to match with the questions might also partly determine the questions possible to answer (Bauman and Nutbeam, 2014). As previously mentioned, the RCT is seen as the gold standard to test the effectiveness of a program. However, when evaluating a policy this is most likely not a feasible design as there might not be a control group (i.e. if implemented nationally), not possible to randomize (policy enacted in a county or municipality) or not enough resources to collect data from the needed number of units (settings, municipalities, counties). Quasi-experimental designs where the control group is not random or where time-series data from existing monitoring are used might be feasible options (Bauman and Nutbeam, 2014). Natural experiments (Craig et al, 2012), and step-wedged designs (Hemmig et al, 2015) could overcome some if these challenges, but still requires that the power is sufficient to detect the changes that the evaluation questions aim to find. Pre-experimental designs with pre and post test on the outcome indicators might therefore be the best design to get preliminary indications of effectiveness, and would allow for collecting data on implementation outcomes, barriers and facilitators as well.

A mixed methods design using both quantitative and qualitative methods, as well as using primary (new) and especially secondary (existing) data as sources is useful for policy evaluation. Using existing data or at least data collection tools with known validity, reliability and responsiveness would save resources (De Vet et al, 2011) and ensure comparability of data as much as possible.

IOM stresses that at community level a logic model approach where indicators for each part of the logic models is developed, and the measurement is focusing on the implementation process and the short-term outcomes is the most important (IOM 2013a, pages 239-241). They further stress the need to use standardized measures if possible to allow for comparison across sites, and that such comparisons across sites - although they might differ in context and details of the programs might be able to show some patterns of what works and what not (IOM 2013a, page 244).

2) Implementation of policies or programs

In order for the policy, or the program and activities resulting from it, to have an effect, it needs to be implemented. Implementation is a process of importance to many different disciplines and thus a plethora of definitions, theories, models and frameworks exists. There has been an increased interest implementation research recently, also within the field of public health.

Theories, models and frameworks for implementation

The Diffusion of innovation theory has been influential in understanding implementation in many disciplines although originating form studying spread of agricultural innovations (Rogers, 2003). The idea that adopters vary in how quickly they adopt a new innovation (idea, tool, policy) indicates that innovators or early adopters who are willing to try out the policy and provide feedback on how well it works are more valuable in trying out a new policy than the late majority who want to see proof before changing their practice. The theory also specifies that innovations need to be compatible with prevailing values, have a relative advantage compared to current practice, be simple and flexible, possible to try out without committing too much and the results should be observable.

Pfadenhauer et al (2017) recently proposed a framework embracing both implementation and context from a public health perspective. According to Pfadenhauer et al (2017), the **i**mplementation process refers to the *"social processes, through which interventions are operationalised in an organisation or community. The implementation process is an active, multistage, iterative and dynamic process that does not usually occur in a linear fashion"*. The later point is recognized by several other and Bertram et al (2015) offer a four staged process model of implementation – moving from exploration of needs and fit, to installation (acquire resources and prepare organization), to initial implementation and finally into full implementation. Implementation is also an actively planned and deliberately initiated effort with the intention to bring a program into policy and practice within a particular setting or sector (Pfadenhauer et al 2017). The setting refers to the specific physical location, in which the program is put into practice, and the setting is interacting with both the implementation process and the context around it. These interactions with context are what makes evaluation of public health program different from controlled experiments since the contexts can not be controlled for, but only accounted for in the interpretation of the evaluation.

According to Pfadenhauer et al (2017), context *"reflects a set of characteristics and circumstances that consist of active and unique factors, within which the implementation is embedded".*

Furthermore, they see it as an "*overarching concept, comprising not only a physical location but also roles, interactions and relationships at multiple levels"* (micro, meso and macro). In total, they identify and define seven domains of the context;

- geographical
- epidemiological
- socio-cultural
- socio-economic
- ethical
- legal
- political

Many factors influencing implementation are found in these contexts and the divide between the ***inner context*** (within the setting) and the ***outer context*** (outside the setting) has also been acknowledged in previous reports on such factors (Aarons et al 2011; Durlak & DuPre 2008).

An implementation theory attempts to "*explain the causal mechanisms of implementation and it is therefore analogous to a programme theory, which attempts to explain the causal mechanisms linking an intervention and its outcomes*" (Pfadenhauer et al 2017). In line with this, the Intervention Mapping protocol for health promotion program development states in step 5 that once the health promotion program has been developed in the four first steps, these steps should be repeated to make an intervention to implement the program by focusing on the factors determining adoption, implementation and maintenance in order to develop strategies to change these (Eldredge et al 2016).

Broadly speaking, implementation agents **"***comprise all individuals and organizations engaged with (i) deciding to implement a given intervention (e.g., funders, administrators) or (ii) implementing this intervention (e.g., providers, advocates, physicians, nurses*)" (Pfadenhauer et al, 2017). However, ***adopters*** are often seen as the gate keepers within the sector or setting who is asked to implement the policy, whereas the ***implementers*** are those that have to change their routines/make the changes in the setting or the sector (Eldredge et al 2016).

In order to develop an implementation intervention, the adopters and implementers must be defined and their barriers and facilitators for adoption and implementation must be found. These factors can be found using theory, evidence from research as well as new research such as observation and interviews. Durlak and DuPre (2008) found 11 common factors influencing implementation across several reviews of health promotion program implementation - funding, a positive work climate, shared-decision making, co-ordination with other agencies, formulation of tasks, leadership, program champions, administrative support, providers' skill proficiency, training, technical assistance. In a recent case study, Muellmann et al (2017) found that factors facilitating adoption, implementation and maintenance of 6 policies implemented in European countries were: active involvement, good communication between coordinating organizations, sufficient training, tailoring of materials, embedding in organizational structure and secure funding. Aligned with this findings, Bertram et al (2015) sort the factors influencing implementation (the drivers) in to three

types; Leadership, Organization and Competency – and points out that implementation strategies need to strengthen these.

Implementation strategies

When the most important factors influencing adoption and implementation have been found, these have to be matched with strategies that can possibly change these factors and thus enable implementation of the policy/program. Proctor et al (2013) have outlined a set of recommendations on how to specify and report on implementation strategies (i.e. the actor, the action, the action target, temporality, dose, implementation outcome and justification). They defined implementation strategies as "*methods or techniques used to enhance the adoption, implementation and sustainability of a clinical program or practice*", and they acknowledged that the complexity of implementation strategies can vary widely and strategies are often used in combinations, but yet they tried to avoid the word intervention to not confuse this with (clinical) interventions aimed at changing the factors influencing the behavior of the target group. Powell et al (2017) followed up with an expert consensus study resulting in 73 implementation strategies for mental health, but Leeman et al (2017) have argued that a better way forward might be to classify implementation strategies according to who the actors are and which factors they are targeting, and they propose the following five classes of implementation strategies; dissemination, implementation process, integration, capacity-building and scale-up.

Implementation outcomes

Finally, the implementation strategies should result in some implementation outcomes. Proctor et al (2011) defined eight implementation outcomes, which can be relevant in different phases of implementation.

Appropriateness, acceptability and feasibility might be particularly important to target and assess in the pre-implementation and the early phases of implementation***. Appropriateness*** is the perceived fit, relevance, or compatibility of the innovation for a given setting and/or the perceived fit of the innovation to address the problem. ***Acceptability*** is similar to appropriateness, but is closer to the actual implementation and is defined as the perception among implementers that an innovation is agreeable, palatable or satisfactory. It is possible to agree to the appropriateness of an innovation, but still not find it acceptable within own practice. This might be much related to perceptions of ***feasibility*** which is defined as the extent to which an innovation can be successfully used or carried out within a given setting. These three outcomes can be measured through questionnaires or interviews. Adoption is also an important outcome early on and a prerequisite for implementation. **Adoption** is defined as the intention, initial decision or action to try or employ an innovation, and can be measured by keeping track of decisions to adopt at the setting or sector level.

Fidelity would require some period of implementation to have passed, as will cost assessment. **Fidelity** is defined as the degree to which an intervention was implemented as it was prescribed in the original protocol or as it was intended by the program developers, and observation is the best way to ascertain this although check lists and interviews are often used. **Cost** is defined as the cost of an implementation effort, and might be important for adoption or maintenance. These are complex

to measure, but within the RE-AIM framework there will be more focus on this moving forward (Glasgow et al, 2019).

Finally, penetration and sustainability are more towards the maintenance phase of implementation. **Penetration** is defined as the integration of a practice within a service setting or its subsystems, which may be relevant when wanting to change the practice of teachers within a school or couches within a sports club, or even supermarkets within a chain**. Sustainability** is defined as the extent to which a newly implemented innovation is maintained or institutionalized within a setting's ongoing, stable operations. This can be assessed through questionnaires or surveys.

Having provided an overview of the two main topics - evaluation and implementation - in this section, the next section of this document will describe a protocol for writing implementation and evaluation plans applying this knowledge.

PROTOCOL for Co-Creating implementation and evaluation plans

An overview of the steps in this protocol is given in table 1. The steps are aimed at guiding the CO-CREATE partners in choosing one to three policy ideas from their youth alliance and developing implementation and evaluation plans for these in collaboration with relevant local stakeholders. Although some of the tasks in the steps are required before the next steps, other tasks might need to be revisited while primarily working on the tasks in a subsequent step. Given the participatory nature of the execution of this protocol, each CO-CREATE partner has great flexibility in the order and how the steps and tasks are done as long as the Implementation and Evaluation plan with it's five parts in the local language is delivered in the end.

There are 5 person months of time provided to each CO-CREATE partner in WP7 for the task of writing the implementation and evaluation plans, but there is no budget for running costs. Thus, the involvement of other stakeholders would be based on voluntary basis and mutual interests. Any cost related to workshops or meetings, would have to come from other sources within or outside CO-CREATE.

Given that the types of policy ideas, the local context supporting the idea and the background knowledge in evaluation, implementation and system science varies greatly across the CO-CREATE partners, this protocol will be executed in the CO-CREATE spirit. This means sharing experiences and solving problems through monthly all country online meetings, monthly (and ad hoc) individual country online meetings and e-mail exchanges, as well as continuously update of the resources (templates, references etc) on the joint CO-CREATE Sharepoint. Webinars and workshops on specific topics and skills will be considered when the profile of the core implementation and evaluation teams are known. Organizing this support is the responsibility of the WP7-leader.

Table 1 Overview of the steps, inputs and outputs for writing implementation and evaluation plans for one-three selected policies per country from the Youth Alliances in CO-CREATE.

STEPS	INPUTS	OUTPUTS
1) Select 1-3 policy ideas	• Policy ideas from youth alliances that have been discussed at dialog forums • Statements for rating the policy ideas • Facilitators, Co-facilitators, youth alliances members	• 1-3 selected policy ideas pr CO-CREATE partner
2) Assemble the core team & advisory committee *(consider the need for a team*	• 1-3 selected policy ideas • Arguments of what is in it for the members not on CO-CREATE funding • A designated Core-team leader from the CO-CREATE partner • Recruited stakeholders/core team members	• A list of members of the core team and the advisory committee • A brief outline of agreements on tasks, roles and responsibilities

per policy idea selected)	• An initial joint meeting of core team and advisory committee to agree on tasks, roles and responsibilities	
3) Draw logic models for policy and for Implementation *(for each selected policy idea separately)*	• System maps and policy forms from youth alliances • Resources on why and how to draw logic models (i.e. the community tool box)/system maps (i.e. group model building) • Worksheets for logic models on policy and implementation (appendix 2 & 3) • Two workshops to jointly make these logic models and collect input for the implementation and evaluation plans through the worksheets • Template for overall evaluation framework (appendix 1)	• (Systems) Logic model for policy • (Systems) Logic model for implementation • Worksheets with input for implementation and evaluation plans • First draft of the overall evaluation framework (appendix 1) based on the logic models and worksheets
4) Write up the implementation plan *(for each selected policy idea separately)*	• Implementation knowledge/expertise • Logic model of implementation & Worksheets for logic model of implementation • Template for implementation plan (appendix 4) • Cycles of drafting, consultations and revisions between the core team leader and the rest of the team and/or advisory committee	• Implementation plan part of D7.6 • Revised overall evaluation framework (appendix 1)
5) Focus the evaluation *(for each selected policy idea separately)*	• Evaluation knowledge/expertise • Logic models, worksheets and implementation plan • Template for evaluation plan (appendix 5) • Cycles of drafting, consultations and revisions between the core team leader and the rest of the team and/or advisory committee	• Evaluation plan part for D7.6 • Finalized overall evaluation framework (appendix 1)
6) Finalizing the Implementation and evaluation plan (D7.6)	• Logic models of policy and of implementation (Step 3) • Combined evaluation framework (Step 3-5) • Implementation plan (Step 4) • Evaluation plan (Step 5)	• A brief report pulling together the 5 parts of the Implementation and evaluation plan for the policy (D7.6)

STEP 1) Select one to three of the policy ideas from the Youth Alliances

After the Youth Alliances have develop their policy ideas (WP5) and discussed and revised these based on the input from relevant stakeholders in the dialogue forums (WP6), a selection of 1-3 of these policy ideas for the implementation and evaluation plans must be done. All policy ideas that have been discussed at the dialog forum and not dismissed by the Youth Alliances afterwards, are eligible for consideration.

Existing policy windows and alignment with existing mandates can increase the likelihood that input from CO-CREATE, including the policy implementation and evaluation plans, can be used by policymakers as public consultation material. Similarly, the institutionalization and possibility to demonstrate short term impact are important for uptake by politicians. However, in the spirit of CO-CREATE the system perspective of addressing feedback loops, the innovativeness and potential for scale-up and also what is most supported by youth/rooted in their perceptions should be taken into consideration in the selection process.

Within the alliances the eligible policy ideas should be considered based on the statements below. Each statement is rated by youth, facilitators and co-facilitators on a five point scale - strongly disagree (1) to strongly agree (5). The ratings for each statement are summed to an overall score for the policy idea. The policy idea with the highest score is the selected policy idea from the Youth Alliance. The scoring can be done in a plenary session if there are few members in the alliance and else discussed and scored in groups first and then scored in a plenary session where each group give their score on each statement and an average is made before summing it up.

Statements for rating:

The policy idea…: (strongly disagree, …… strongly agree)

- reflects the youths view of the problem and solution
- is innovative
- is taking a systems perspective (addressing a feedback loop(s) or levers in the obesity system)
- is likely to have the support of the key stakeholders
- is likely to be enacted in the near future
- is aligned with relevant policies, strategies or frameworks being mandated by municipal governments
- is likely to be institutionalized and maintained over time

If the core implementation and evaluation team (see next STEP) does not have capacity to make implantation and evaluation plans for three policy ideas, they will repeat the scoring for the three top ideas from the youth alliances in order to prioritize among these three ideas.

STEP 2) Assemble the core team and an advisory committee

Throughout the policy cycle, the stakeholders should always be engaged and included in communication and activities. However, for the core implementation and evaluation team this would require a bit more commitment and thus it needs to be clear why each one is invited, what contribution will be expected from each one and how will each one be rewarded for his/her contribution. The kind of rewards are likely to differ by country, thus the CO-CREATE team will share, discuss and compile an overview of what these might be for the various stakeholders involved including the youth.

- ***Recruit the leader(s) of the team***
    - This should be a person from the CO-CREATE partner (preferably one that has attended the work shop (D7.5)) in order to ensure that the deliverable for CO-CREATE is completed and that youth are properly involved.
    - If possible, a stakeholder with ownership to the policy implementation and evaluation process should co-lead the core team. This could be a person from the municipality or setting in which the policy will be implemented, or a Non-Governmental Organizations (NGO) that might be the support system during the implementation.
    - It could also be considered to have a youth from one of the alliances as part of leadership trio in order to ensure their views are always present.
    - The main task of the leader(s) of the team is to lead the process outlined in this protocol and deliver 1-3 implementation and evaluation plans.
- ***Recruit the rest of the members of the team***
    - These members are involved through the whole process of writing the plans
    - These members should cover the following stakeholders:
        - A youth representative (preferably representing the alliance proposing the idea, but it depends on interest and availability),
        - A stakeholder who would be involved in adopting and implementing the policy (representing the setting),
        - A stakeholder providing support of the implementation (i.e. NGO or public bodies from the municipality).
    - If not already covered by the leaders or the members above, the following expertise should preferably be included from the CO-CREATE partner or else through external collaboration:
        - A person with expertise in evaluation
        - A person with expertise in policy implementation
        - A person with expertise in the subject matter (diet, physical activity etc).
- ***Establish an Advisory committee (optional)***
    - If there are stakeholders or experts that are not able to commit to being on the team, these might be willing to be in an advisory committee.

- Think of special interest groups, policy makers, representatives from the sector or settings that will implement the policy, others with an investment in the policy and its outcomes, or the general public.

- ***Convene the team (and advisory committee) to a first meeting (approx. 2 hours)***
  - Present the CO-CREATE project and the task of writing implementation and evaluation plans to them briefly
  - Discuss and agree upon the following (summarize the conclusions in the minutes):
    - needs for capacity building in system science, implementation and/or evaluation
    - roles and responsibilities
    - tasks and timelines for the three following steps
  - Present the one-three policy ideas from the youth alliances, and if needed have the core team and advisory committee rate them according to the statements in order to narrow down to one (or two) of these ideas to work with moving forward.
  - Summarize the process and the conclusions in the minutes.

STEP 3) Draw logic models for the policy and for implementation

The Centre TRT's evaluation framework – see Appendix 1 - provides an easy to grasp overview of the evaluation and implementation (Leeman at al, 2012), but in order to get to such overviews agreement on the logic of change for both the policy and the implementation is needed. Drawing logic or system models for these processes is recommended. The community tool box provides background and examples of how to draw logic models (see Chapter 2, Section 1 of the online Community Tool Box - https://ctb.ku.edu/en), whereas Hargreaves (2010) provides guidance on how to apply a systems lens on the evaluation. When drawing the models, attention should also be paid to think of potential unintended consequences of the policy and its implementation, as well as whether there might be persons or groups likely to be opposed to the policy.

For the logic model/system map of the policy idea, the core team should refine the map drawn in the youth alliances, whereas for the implementation logic model the core team has to start from scratch. The leader should make a first draft

Once the logic models are in place, the core team can use the templates in Appendix 2 and 3 to distill the processes that are most important and articulate the core team's joint understanding of the *inputs* and *activities* going into the policy and implementation logic of change, as well as the *outputs* and *outcomes* to which the inputs and activities should lead.

Organize two workshops with the core team in order to draft and refine the logic models for the policy and the implementation.

- ***Workshop 1 (approx. 3 hours)***
  - Hour 1
    - Provide the needed training in systems thinking, the purpose of logic models and how to make them.
  - Hour 2
    - Provide them with the system map from the Youth Alliance in order to use this as an input for their own logic model for the policy idea
    - When the logic model for the policy has been drawn, provide them with worksheets for logic models of the policy (appendix 2) and ask them to start filling it in.
    - Write down questions/issues to collect further information on, and assign it as tasks (by whom and by when).
  - Hour 3
    - Draw logic models for the adoption and implementation of the policy (one for each).
      - What facilitating factors might make adoption and/ or implementation of the policy more likely?
      - What barriers might make adoption and/or implementation of the policy less likely?
    - When the logic model for the implementation has been drawn, provide them with worksheets for logic models of the implementation (appendix 3) and ask them to start filling it in.
    - Write down questions/issues to collect further information on, and assign it as tasks (by whom and by when).

- **Between the workshops**
  - The leader(s) of the team refines the logic models and the worksheets based on inputs from the team members on the questions and consultations with the advisory committee and/or the CO-CREATE team on this task.
  - The logic model of the factors influencing adoption/implementation should be checked against theories and existing evidence.
  - Based on the above refined documents, the leader starts filling in the Centre TRT evaluation framework (Appendix 1).
    - ***Inputs***
      - The problems and solutions are preliminary defined by the youth alliances (policy form, WP5) and stakeholders in the dialog forum (refined policy idea, WP6).
      - The politics of the implementation process and other factors related to resources and monitoring will need to be discussed and described by the team.

- ***Activities***
  - The *formulation* of policy is done by the Youth Alliance (WP5) with input from stakeholders in the dialogue forums (WP6), and this will be used although there is likely to be changes to this in an enactment process or even as part of writing these plans.
  - The *enactment* of the CO-CREATE policy ideas are likely to still be hypothetical when drafting the implementation and evaluation plans, but consideration of support and enactment is one of the statements rated when selecting policy ideas (see step 1).
  - The *implementation* of the policy requires a good understanding of which activities are needed to implement the policy/local programs and how these activities will overcome barriers or strengthen/install facilitating factors (See Step 4).
  - The *maintenance/modification* of the policy/local program requires the core team to think beyond the initial implementation phase to ensure that the policy/local program is overturned by new political process and that it will have sustained effects. This may involve securing long term funding/resources, as well as policy/program modifications based on the results of executing the final implementation and evaluation plans.
- ***Outputs of the…***
  - ..formulation activities will be the policy ideas from the Youth Alliances after the dialog forums.
  - ..enactment activity would be hypothetical (a decision made by the body in power to make a decision to enact the policy).
  - ..implementation activities are the key indicators that the implementation activities have been followed, i.e. adoption by settings and sectors, implementation according to plan in each setting/sector, reaching the intended target population, funding secured.
  - ..maintenance activities would describe any modifications made, as well as agreement of further funding, partnerships, implementation.
- ***Outcomes*** are the desired as well as unanticipated results of a policy (i.e. the effectiveness evaluation). The desired outcomes are divided in short term (environment), intermediate (social norms, behaviors) and long term outcomes. Long term outcomes focus on the public health impact of achieving the population level health outcomes and ensuring that these contribute to equitable distribution as well as the improvements being cost-effective. Unanticipated results could occur as the policy implementation will be part of a system which responds in other ways than what the policy was planned for, and foreseeing and monitoring this is part of the outcomes as well. Some of these outcomes can be measured through leveraging existing

surveillance or monitoring systems, while others will require new data collection (and additional funding).

- The intended changes in environment, behaviors of the target group and in social norms should be found in the logic model/system map of the policy idea made by the core team in workshop 1.
- The logic model/system map could also be a source for ideas of unanticipated results of the policy.
- The long term impacts would be on population rates of overweight and obesity among youth and ensuring that any pre-existing social inequalities (socio-economic, gender, ethnicity, geographically) in these rates are evened out.
- Assessing cost-effectiveness would require having an economist in the team/the advisory committee that could be drawn on for the evaluation plan of this.

- ***Workshop 2 (approx. 1.5 hours)***
  - Present the refined logic models, worksheets and the draft Centre TRT evaluation framework
  - Ensure that the inputs received are correctly incorporated
  - Discuss and refine the logic model for both the policy and it's implementation

STEP 4) Write up the implementation plan

The implementation activities/strategies in the implementation plan should be carried out by a support system (outside of the setting) and a delivery system (inside the setting). The composition of the team writing the implementation plan should reflect this, which might require additional support from the advisory committee, strengthening of the expertise in implementation and/or a member from an NGO or public body that is willing to be the support system. There should also be at least one youth representing the type of setting where the policy should be implemented.

The implementation activities/strategies should be designed to address the factors that determinate adoption and implementation of the policy as specified by the logic models of adoption and implementation in Step 3. Step 3 has also provided ideas on inputs on the activities through the worksheet.

The process of writing up the implementation plan consists of cycles of drafting, consultations and revisions. The team leader(s) has the responsibility for the drafting, but consults with the core team, the advisory committee and potentially new adopters and implementers with a fresh eye on the compatibility and feasibility of the implementation activities. Consultations are done by joint meetings, e-mails (to the team/committee or individually), phone conversations and if needed on site visits at potential settings for implementation.

- **Draft the implementation plan** by finding answers to the following questions:
  - Which organization will be the support system?
    - What are their resources and capabilities?
  - In which settings will the policy be implemented?
  - Who are the adopters in these settings and what are the procedures of adopting new policies?
    - Through which channels are the adopters most likely to be reached? What kind of messages will make them listen?
    - What activities would enhance the facilitating factors and which would diminish the barriers to adoption?
  - Who are the implementers?
    - What activities would enhance the facilitating factors and which would diminish the barriers?
- **Check the answers to the questions by:**
  - Reviewing the literature on implementation strategies/activities
  - Checking the feasibility of the proposed activities with the support system
  - Checking the feasibility of the proposed activities with the adopters and implementers

At the end of STEP 4, the outputs are a draft implementation plan (Appendix 4) and an updated overall evaluation framework (Appendix 1) for D7.6.

STEP 5) Focus the evaluation

Based on the implementation plan (step 4) and the overall evaluation framework (step 3/4), the team now has to focus the evaluation and write up the evaluation plan. The Center TRT provides Evaluation Framework Questions which should be consulted when making these decisions (http://centertrt.org/?p=evaluation_questions). The output of this step is the Evaluation plan for D7.6.

At this step the team needs to make sure to have an expert on evaluation, preferably one who would be involved with the actual evaluation if the policy is enacted. This person could be a researcher from the CO-CREATE partner or an employee of the organization that will support implementation of the policy (i.e. the municipality and/or an NGO).

The formative evaluation where the rationale for the policy, its theory of change as well as implementation activities are outlined, is taken care of by the previous steps. Thus, the plan is for the process evaluation of the implementation and an outcome evaluation and the purpose of the evaluation is to provide useful recommendations for decision makers about the implementation and effects of the policy. If there is a need for short term evidence in order to secure continued support and funding, process evaluation might be more useful than outcome evaluation. If the policy is likely to be implemented for a longer term, a combined process and outcome evaluation is better.

The decision of what to evaluate should be based on what the team judge to be feasible and most beneficial to the policy idea based on the assumptions made in the previous steps. Feasibility would need to take into consideration the likely resources (monetary and human), the available expertise, the amount of time/timing of the evaluation, any ethical considerations with regards to data collection, the level of rigor and accuracy required of the results and whether it would be a one-time data collection or multiple measurement points, and whether to collect new data or use of existing surveillance systems.

The process of writing up the evaluation plan consists of cycles of drafting, consultations and revisions. The team leader(s) has the responsibility for the drafting, but consults with the core team and the advisory committee. Consultations are done by joint meetings, e-mails (to the team/committee or individually), phone conversations and if needed on site visits at potential settings for implementation. It is advised to take a meeting with the team at the beginning to clarify the **Users and uses** and the **Evaluation questions**, whereas the technical **methodological choices** of design, data collection methods and indicators can be done by a smaller working group and presented to the core for discussion and agreement in a second meeting.

- **Users and uses -** Think about who this evaluation is likely to be done for and what would they like to use the results for?
    - Is it done for the funders to demonstrate some first promising outputs like adoption, implementation and reach? Is it done for the implementers in order to improve the implementation? Is it done for the public health authorities to show effect on environments, behaviors of the target group, social norms within an intermediate time frame? Or in social inequalities in overweight/obesity in the long term?
- **Evaluation questions –** Specify the evaluation questions, the participants and boundaries of the evaluation (one or multiple settings?, One or multiple municipalities?, will there be a control group?, will both adopters/implementers and the target group provide data to answer the questions?)
    - Are the questions about:
        - Implementation processes: Adoption? Reach? Degree of implementation according to plan? Perceptions of the implementation process from the view of adopters, implementers and/or target group?
        - Effectiveness: Changes in environments (which?)? Changes in behaviors (which?)? Changes in social norms? Changes in (social inequalities in) overweight/obesity? Costs of the policy?
- **Methods** - Specify the research design, data collection methods and indicators to answer the evaluation questions, as well as ethical issues related to the data collection.
    - Research design and power
        - Quasi-experimental designs where the control group is not random or where time-series data from existing monitoring are used might be feasible options. Pre-experimental designs with pre and post test on the short term outcome indicators might be the best design to get preliminary indications of effectiveness.

- Process evaluations could be formative or summative, where the former will feed the data back into the implementation process in order to improve it and the latter is written up after all the data are collected and analysed.

- Data collection methods
  - Data can be collected through quantitative and qualitative methods. Quantitative data are typically collected through measurement instruments (I.e. scales, measurement tapes, questionnaires), counts through direct observation, sales data etc, whereas qualitative data include case studies, interviews, observations, content analysis of documents etc.
  - Data could be collected as primary data to answer the evaluation questions for this policy. This would require finding and assessing the fit and quality (validity, reliability, responsiveness) of existing tools or developing and testing the quality of new tools (De Vet et al, 2011). Finding or developing measurement tools is a time consuming process, thus using data that are already being collected (secondary data) should always be considered. Contributing to establishing standards for how to measure the short term outcomes or changes in environment may be considered if resources allow.
- Indicators
  - Are specific, observable, measurable characteristics of changes that demonstrate progress toward outcome or impact.
  - Using the Logic Model Worksheets in Appendix 2 and 3 enables the evaluation team to note indicators for each part of the Evaluation Framework while mapping it out.
  - Existing measurement tools that might be usefull:
    - Table H-3 (page 430) in Appendix H of the IOM report (2013) lists some resources on tools for evaluating community obesity prevention initiatives
    - Some of the evaluation plans on the Center TRT website also provide references to measurement tools (http://centertrt.org)
    - For implementation outcomes and context there are several resources i.e.
      - Measures for implementation research (https://www.c4tbh.org/resources/measures-for-implementation-studies/)
      - RE-AIM (http://www.re-aim.org/)
      - National Cancer Institute GEM-Dissemination and Implementation Initiative (GEM-D&I) (https://www.gem-measures.org/public/wsoverview.aspx?cat=8&aid=0&wid=11)

- Ethics
  - The team has to consider the ethical aspects of the data collection methods and measures chosen, and specify how ethical clearance and procedures for handling personal data should be handled as per required procedures in each country.

STEP 6) Finalizing the Implementation and evaluation plan

Based on steps 3-5, the team has produced the following five outputs that should be combined in a final report for each policy which would then be the Implementation and evaluation plan (D7.6):

- Logic models of policy (Step 3)
- Logic model of implementation (Step 3)
- Combined evaluation framework (Step 3-5)
- Implementation plan (Step 4)
- Evaluation plan (Step 5)

It is the responsibility of the team leader(s) to write this report and have it approved by the core team.

References

- Aarons GA, et al. Advancing a Conceptual Model of Evidence-Based Practice Implementation in Public Service Sectors. Adm Policy Ment Health 2011; 38**,** 4–23.
- Bauman A, Nutbeam D. *Evaluation in a nutshell*. (2nd ed). Australia, McGraw Hill Education, 2014.
- Bertram RM, et al. Improving Programs and Outcomes: Implementation Frameworks and Organization Change. Research on Social Work Practice 2015; 25: 477-487.
- Eldredge LKB, et al*. Planning Health Promotion Programs: An Intervention Mapping Approach* (4th ed). USA: Jossey-Bass, 2016.
- Centers for Disease Control and Prevention. *Overview of CDC's Policy Process*. Atlanta, GA: Centers for Disease Control and Prevention, US Department of Health and Human Services; 2012.
- Centers for Disease Control and Prevention. *Using Evaluation to Inform CDC's Policy Process*. Atlanta, GA: Centers for Disease Control and Prevention, US Department of Health and Human Services; 2014.
- Craig P, Cooper C, Gunnell D*, et al* Using natural experiments to evaluate population health interventions: new Medical Research Council guidance. J Epidemiol Community Health 2012; 66: 1182-1186.
- De Vet HCV et al. *Measurement in medicine. A practical guide*. Cambridge University Press, 2011.
- Durlak JA, DuPre EP. Implementation Matters: A Review of Research on the Influence of Implementation on Program Outcomes and the Factors Affecting Implementation*.* Am J Community Psychol 2008; 41: 327-350.
- Glasgow RU, et al. RE-AIM planning and evaluation framework: adapting to new science and practice with a 20 -year review. Front Public Health 2019; 7:64.
- Hargreaves MB. *Evaluating system change: A planning guide*. Mathematica Policy Research Inc., 2010.
- Hemming K, Haines TP, Girling AJ, Lilford RJ. The stepped wedge cluster randomised trial: rationale, design, analysis, and reporting. BMJ 2015; 350: h391.
- Horodyska K, et al. Implementation conditions for diet and physical activity interventions and policies: an umbrella review. BMC Public Health 2015, 15.
- Institute of Medicine 2012. *Accelerating progress in obesity prevention: Solving the weight of the nation.* Washington, DC: The National Academies Press.
- Institute of Medicine 2013. *Evaluating obesity prevention efforts: A plan for measuring progress*. Washington, DC: The National Academies Press.
- Leeman J, et al. An Evaluation Framework for Obesity Prevention Policy Interventions. Prev Chronic Dis 2012; 9: 110322.
- Leeman J, et al. Beyond "implementation strategies": classifying the full range of strategies used in implementation science and practice. Implementation Sci 2017; 12: 125.

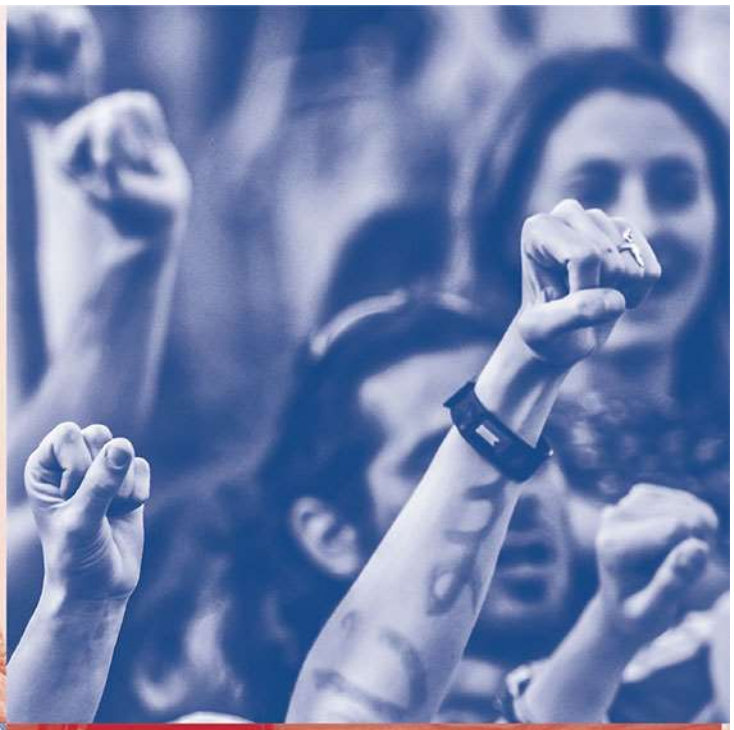
- Muellmann S, et al. Views of policy makers and health promotion professionals on factors facilitating implementation and maintenance of interventions and policies promoting physical activity and healthy eating: results of the DEDPIAC project. BMC Public Health 2017; 17: 932.
- Nilsen P. Making sense of implementation theories, models and frameworks. Implementation Sci, 2015; 10, 53.
- Pfadenhauer LM et al. Making sense of complexity in context and implementation: the Context and Implementation of Complex Interventions (CICI) framework. Implementation Sci 2017; 12**:** 21.
- Powell BJ, et al. A refined compilation of implementation strategies: results from the Expert Recommendations for Implementing Change (ERIC) project. Implementation Sci 2015; 10: 21.
- Proctor E, et al. Outcomes for implementation research: Conceptual distinctions, measurement challenges, and research agenda. Adm Policy Ment Health 2011; 38: 65-76.
- Proctor EK, et al. Implementation strategies: recommendations for specifying and reporting. Implementation Sci 2013; 8**:** 139.
- Roberto CA et al. Patchy progress on obesity prevention: emerging examples, entrenched barriers, and new thinking. Lancet 2015; 385: 2400-2409.
- Rogers EM. *Diffusion of innovations.* (5th ed). New York, Free Press 2003.
- Saunders RP, et al. Developing a Process-Evaluation Plan for Assessing Health Promotion Program Implementation: A How-To Guide. Health Promotion Practice 2005; 6: 13
- World Health Organization. *WHO global strategy on diet, physical activity and health: a framework to monitor and evaluate implementation*. 2008.
- World Health Organization. *WHO evaluation practice handbook*. 2013.
- Zukoski A, Luluquisen M. Participatory evaluation. What is it? Why do it? What are the challenges? *Community-based Public Health Policy and Practice*, 2002, No. 5

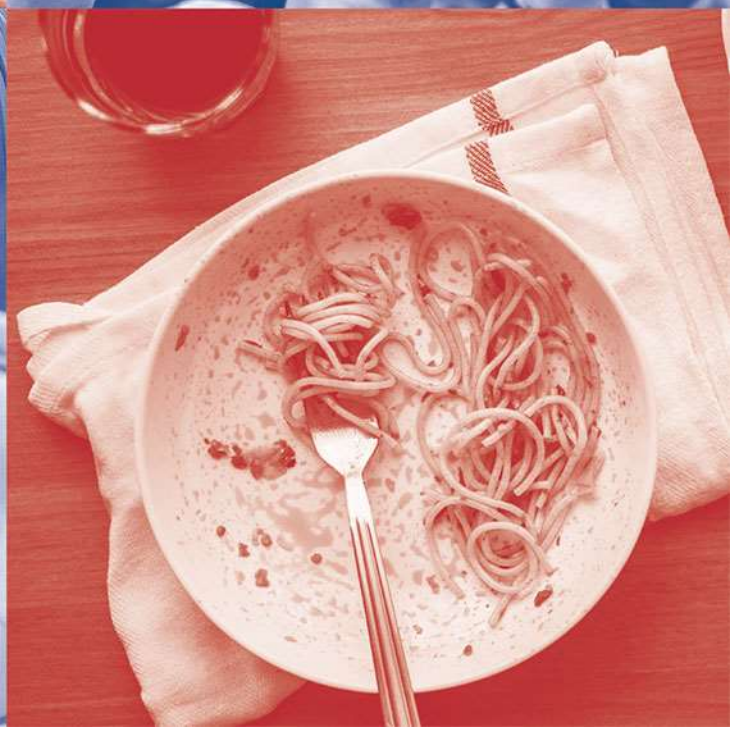
Appendix

1. Word template for the overall Center TRT's evaluation framework
2. Template Logic Model Worksheet for the policy (based on the Evaluating Violence and injury prevention policies briefs)
3. Template Logic Model Worksheet for adoption and implementation
4. Template for writing implementation plans
5. Template for writing evaluation plans (based on the Center TRT's evaluation plan examples)

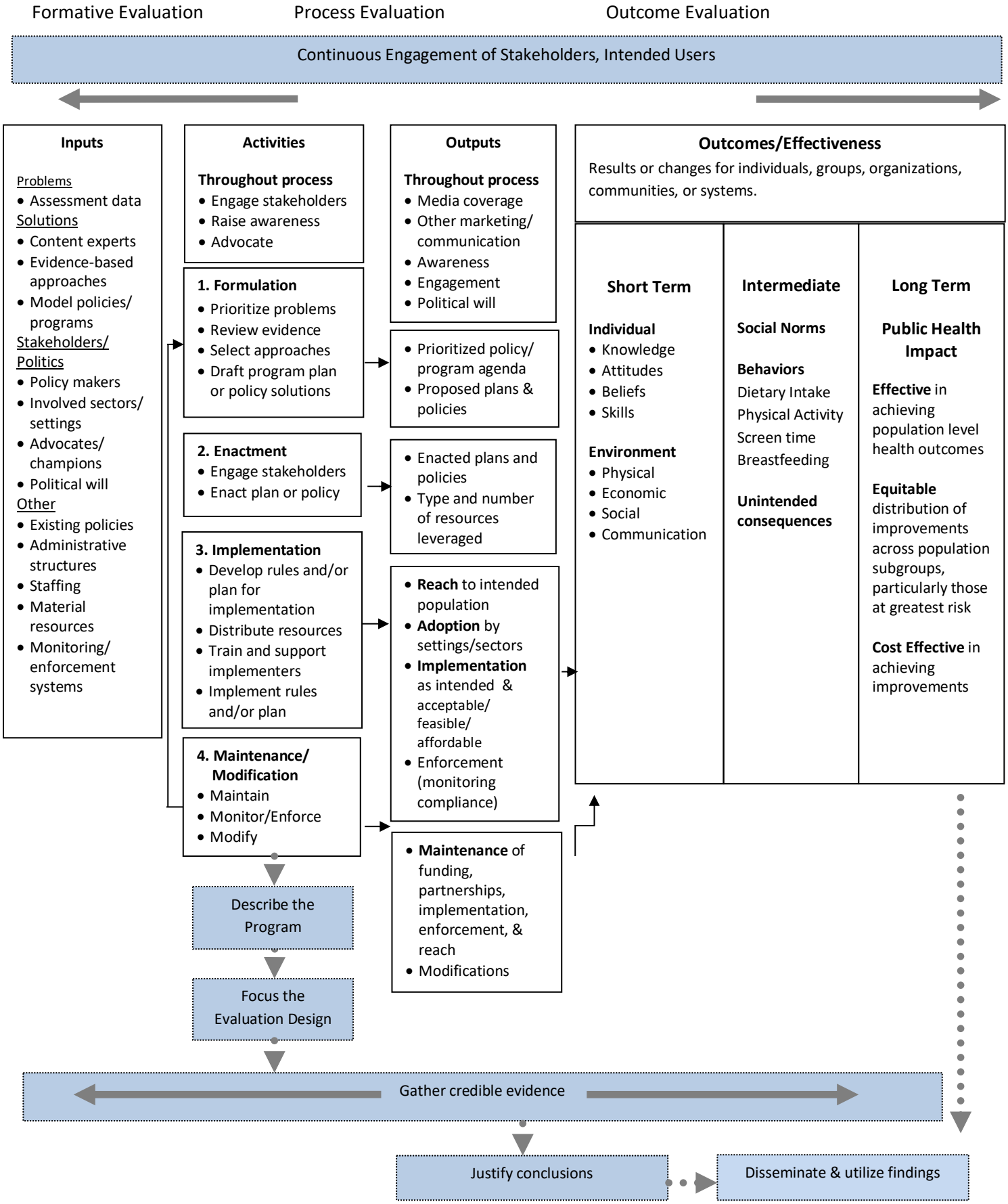
The CO-CREATE project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 774210. The products of the research are the responsibility of the authors: the European Commission is not responsible for any use that may be made of them.

CO-CREATE





Center TRT Evaluation Framework



Shaded portions represent the CDC evaluation framework. 07.24.12

Appendix 2 Logic model worksheet for the Policy idea (based on Appendix H of the CDC Evaluating Violence and Injury Prevention Policies Briefs, https://www.cdc.gov/injury/pdfs/policy/Appendices-a.pdf)

Context/need:	**Assumptions /Theory of Change (logic model)**			**External influences** *(Other contextual factors which could influence these outcomes - systems model. Any potential unintended consequences?)*		
	Inputs	**Activities**	**Outputs**	**Short-term outcomes**	**Intermediate outcomes**	**Impacts**
	Indicators	**Indicators**	**Indicators**	**Indicators**	**Indicators**	**Indicators**

Appendix 3 Logic model worksheet for Adopters and Implementers

a) **Adopter(s)** *Fill in one worksheet per type of adopter*

Setting: **Context/need:** **Adopter:** **Procedures for adoption:**	**Assumptions /Theory of Change (logic model)** *(Barriers and facilitators of adoption in the inner and outer context)*			**External influences** *(Other contextual factors which could influence these outcomes - systems model. Any potential unintended consequences?)*		
	Inputs	**Activities**	**Outputs**	**Short-term outcomes**	**Intermediate outcomes**	**Impacts**
	Support team	**Dissemination strategies** *(messages/materials & distribution of it)*	Number of settings reached and responses	*Adopters express awareness and positive attitudes*	*Policy adopted*	
	Support team	**Integration strategies**	*Resources allocated* *Implementation team set up* *Plan for implementation agreed upon*	*Reduced internal barriers to implementation* *Leadership/organization supportive of implementation*		
	Indicators	**Indicators**	**Indicators**	**Indicators**	**Indicators**	**Indicators**
			<u>Reach</u> *Number and types of settings contacted versus number/representativeness of positive responses*			

b) **Implementer(s)** *Fill in one worksheet per type of implementer*

Setting: **Context/need:** **Implementer:**	**Assumptions /Theory of Change (logic model)** *(Barriers and facilitators of implementation at the personal and inner context)*			**External influences** *(Other contextual factors which could influence these outcomes - systems model. Any potential unintended consequences?)*		
	Inputs	**Activities**	**Outputs**	**Short-term outcomes**	**Intermediate outcomes**	**Impact**
	Support team **Delivery system/ implementation team**	**Capacity-building strategies**	Implementers trained Support system adapted to further needs	*Implementers express awareness, positive attitudes and self-efficacy*	Implementers provide feedback on feasibility	Policy adapted and implemented with a sufficiently high fidelity
	Indicators	**Indicators**	**Indicators**	**Indicators**	**Indicators**	**Indicators**

Appendix 4 Template for writing implementation plans

IMPLEMENTATION PLAN ***Name of policy***			
Implementation support system members and responsibilities: **Purpose:** **Setting:** **Adopters and implementers:**			
ADOPTION			
Activities/Implementation strategies	**Targeting which factors?**	**Who are doing this?**	**When and where?**
Activity 1			
Activity 2			
Activity n			
IMPLEMENTATION			
Activities/Implementation strategies	**Targeting which factors?**	**Who are doing this?**	**When and where?**
Activity 1			
Activity 2			
Activity n			

Appendix 5 Template for writing evaluation plans (based on the Center TRT's evaluation plan examples http://centertrt.org/)

EVALUATION PLAN ***Name of policy***		
Evaluation team members and responsibility: **Purpose:** **Evaluation question:** **Ethics/data handling approval procedure:** **Design:** **Data collection** *(incl. available measurement tools)***:**		
PROCESS EVALUATION		
Evaluation questions	**Data to be collected**	**Data collection method**
Reach		
Adoption		
Implementation		
OUTCOME EVALUATION		
Evaluation questions	**Data to be collected**	**Data collection method**
Outcome 1		
Outcome 1		
Outcome n		