



Vaasan yliopisto  
UNIVERSITY OF VAASA

Tiphaine Delecour

# **Utilizing program evaluation processes for strategic learning in nonprofits**

Cases of public and volunteer organizations

School of Management  
Master's thesis in Strategic  
Business Development

Vaasa 2024

---

**UNIVERSITY OF VAASA****School of Management**

<b>Author:</b>	Tiphaine Delecour
<b>Title of the Thesis:</b>	Utilizing program evaluation processes for strategic learning in nonprofits: Cases of public and volunteer organizations.
<b>Degree:</b>	Master of Science in Economics and Business Administration
<b>Programme:</b>	Strategic Business Development
<b>Supervisor:</b>	Rodrigo Rabetino
<b>Year:</b>	2024 <b>Sivumäärä:</b> 99

---

**ABSTRACT:**

Strategic evaluation is a tool that organizations can use to evaluate and adapt their strategies, and various tools exist in the business world to do so. Yet, most of the tools that are created deal with the financial perspective as an outcome, leaving nonprofit organizations forgotten as they cannot fully use the existing tools to evaluate their programs, knowing that financial performance is, for them, considered an output. The existing framework within the realm of program evaluation presents evaluation as a means to achieve results that other strategic learning methods cannot; yet, efforts from social organizations to utilize evaluation results towards learning and adaptation of their program often fail. This study investigates how nonprofit organizations could use program evaluation processes to promote strategic learning. The study will explore two views on program evaluation: the theory of change and the utilization-focused evaluation approach.

The study constructed a tentative conceptual framework based on primary and secondary data sources. The initial framework was designed with the help of existing theoretical research and materials, while the final framework was completed thanks to an empirical study with four participating case organizations. The final framework outlines the general program evaluation processes and strategic learning opportunities nonprofits have when evaluating their programs. The case organizations participating in this study are Bloom Finland, the Multicultural Association of Porvoo, Moni-info Seinäjoki, and International House Oulu. The primary data for the empirical study was gathered with semi-structured interviews with eight decision-makers or presiding board members in nonprofit organizations. The study focused on nonprofit organizations within the specific environment of working with internationals in Finland. The study provided controlled variance between public and volunteer organizations as participants.

The findings answered the research question and demonstrated that program evaluation processes can be used for strategic learning in the nonprofit context, although the results showed that adaptations are needed. Altogether, the study recognized eight general program evaluation processes, each linked to their subfactors identified as specifics of the nonprofit organizations. The framework interpreted two theoretical perspectives into one conceptual representation and resulted in a tentative framework to provide evaluators and nonprofits with an evaluation method promoting using the results for strategic learning. The framework remains tentative as real-life tests were not part of this study. The study's conclusion is the framework that can be used to observe whether the program is running as intended and if conditions, actions, users, and uses are appropriately matched.

---

**KEYWORDS:** Strategic learning, theory of change, program evaluation, nonprofits, utilization-focused.

## Contents

1	Introduction	6
1.1	Motivation for the study	6
1.2	Research problem	7
1.3	Research contributions	8
1.4	Thesis structure	9
2	Theoretical background	10
2.1	Understanding strategic learning processes	10
2.2	Understanding program evaluation	14
2.2.1	Using program evaluation	18
2.2.2	Users of program evaluation	20
2.3	Theory of change	22
2.4	Utilization-focused evaluation	26
2.5	Specifics of public and volunteer organizations	30
2.6	Proposed conceptual framework	32
3	Methodology	34
3.1	Research approach	34
3.2	The case organizations	37
3.3	Data collection	40
3.4	Data analysis	41
4	Findings	47
4.1	View and experience on evaluation	47
4.2	Program evaluation processes	49
4.2.1	The choice of evaluator	49
4.2.2	Defining a program	52
4.2.3	Conditions and actions	53
4.2.4	Context	60
4.2.5	Users and uses	67
4.3	Strategic learning	75

4.4	Summary of the results and revised framework	78
5	Discussion	81
5.1	Theoretical implications	81
5.2	Practical implications	83
5.3	Limitations	85
5.4	Suggestions for future research	85
6	References	87
	Appendices	97
	Appendix 1. Interview Questions	97
	Appendix 2. Interviewees' summary	99

## Figures

Figure 1. The exploring strategy framework (Whittington et al., 2023, p. 13)	10
Figure 2. Strategic process and strategy analysis (Blair, 1998; Hastings, 1996; Maleka, 2014; Williams, 2014)	12
Figure 3. Common broken link in strategic thinking (Williams, 2014)	13
Figure 4. The program evaluation umbrella (Giancola, 2020a)	16
Figure 5. Logic model as embedded theory of change. (DuBow & Litzler, 2018)	25
Figure 6. Using the utilization-focused evaluation processes (Ramirez & Brodhead, 2013)	28
Figure 7. Initial framework	33
Figure 8. Research onion for the study (Saunders et al., 2023)	34
Figure 9. Approach to theory development for the study (Saunders et al., 2023)	36
Figure 10. Full data structure of the study based on Gioia's methodology (Gioia et al., 2013)	45
Figure 11. Gioia's methodology model adapted to the study (Gioia et al., 2013)	46
Figure 12. Summary of program evaluation processes	79
Figure 13. Revised conceptual framework	80

## Abbreviations

<b>Theory of Change</b>	<b>ToC</b>
<b>Utilization-Focused Evaluation</b>	<b>U-FE</b>
<b>Multicultural Association of Porvoo</b>	<b>MAPorvoo</b>
<b>Moni-Info Seinäjoki</b>	<b>Moni</b>
<b>International House Oulu</b>	<b>IHO</b>
<b>In other words (Id est)</b>	<b>i.e.</b>

# 1 Introduction

## 1.1 Motivation for the study

Strategic evaluation is one of the steps within the strategic planning process; it allows companies to carefully evaluate the strengths and weaknesses of their strategy (Nilsson et al., 2016) and adapt their actions if needed (Starbuck et al., 2008). Practicing strategic management is assumed to lead to better financial and non-financial performance (Sassia et al., 2019). As strategic planning is still considered one of organizations' most used management tools (Wolf & Floyd, 2017), it is undeniable that the field of evaluation and strategy has been widely researched (Nilsson et al., 2016). Publications, recent and not, gather countless evaluation methods and techniques accompanied by theoretical perspectives (Newcomer & Brass, 2016); yet, it is noticeable that most existing frameworks grant the financial perspective a strong hand. However, when it comes to nonprofit and social programs, the expectations resulting from the evaluation are not based on the financials or revenues produced but on identifying the outcomes and the extent to which the actions helped the customer (Bellamy & Evans, 2017). In recent years, few publications have focused on these core differences, but those who mention them outline theories and attempt to create initial frameworks to compensate for the deficit in program evaluation (Blamey & Mackenzie, 2007). The theory of change or realistic evaluations (Blamey & Mackenzie, 2007), utilization-focused evaluation approach (Patton & Campbell-Patton, 2023), or an initial program evaluation framework have appeared (Newcomer & Brass, 2016) to fill in the gap.

Focusing this research solely on the nonprofit sector is explained by the sector's high need for strategic capabilities despite the organizations' limited capacities and resources (Rudov et al., 2017). Furthermore, as nonprofit programs' leaders struggle to understand whether their organization is making a difference (Benjamin et al., 2023), increasing the available research in the field would provide more content for nonprofit organizations' decision-makers to review and analyze when decisions are necessary. Research demonstrates that when discussed concerning the nonprofit sector, program evaluation may

express the successes or the failures of the social programs but fails to understand how either outcome was reached (Blamey & Mackenzie, 2007), once again justifying the need for future research with a clear focus in the field.

## **1.2 Research problem**

Amidst the existing research and publications on strategy and evaluation listed throughout the years, it is recognized that there is little mention of program evaluation theory or practice (Newcomer & Brass, 2016), although there is an increasing demand for nonprofits to demonstrate the effectiveness of their program (Bellamy & Evans, 2017; Carman, 2010). Consequently, this thesis will answer the call of Williams (2014), who expresses that “although many social organizations espouse learning and adaptation as part of their culture, actual efforts to use evaluation in this way often fall flat” (Williams, 2014, p. 1). This statement is confirmed by others in the field, who demonstrated that evaluations, although often conducted to reach a goal, unfortunately, remain unused (Patton, 2008a). There is a need to complete the strategic learning loop utilizing program evaluation processes in the specific context of nonprofit organizations.

The purpose of this study is to tap into this research opportunity by answering the research question:

How can program evaluation processes be used for strategic learning in nonprofits?

More specifically, the study aims to answer the question from two perspectives. The lens of the theory of change (ToC) (Weiss, 1998), a theory-based evaluation approach recognized as a reactive and relevant evaluation method to ensure the utilization of the results (Christie & Alkin, 2003); and the lens of utilization-focused evaluation (U-FE) identified as one of the most promising evaluation approaches in terms of use and development (Stufflebeam, 2001), and with a core focus on delivering evaluation for intended users and uses (Patton & Campbell-Patton, 2021a). The goal is to provide the nonprofit sector

with a final framework that can be used to guide evaluations. The aforementioned is achieved by researching existing literature on program evaluation and both views to outline the basis for the framework. The initial framework constitutes the study's starting point, encompassing four case nonprofits public and volunteer organizations, chosen for their methodological fit (Edmondson & McManus, 2007). The case nonprofits are Bloom Finland, the Multicultural Association of Porvoo (MAPorvoo), Moni-info Seinäjoki (Moni), and International House Oulu (IHO). The research used a qualitative method to gather complex social processes linked to the research question (Eisenhardt & Graebner, 2007). The data was collected through semi-structured interviews, which constitute the primary data. Secondary data, in the mean of existing sources, was also used. The study was explored from theory to findings to create the final framework and ensure that the tool is theoretically acceptable, then tailored to the given context, meaning the nonprofit sector.

### **1.3 Research contributions**

The contribution of this study is threefold. Firstly, the study will contribute to the literature in two ways. The final framework constitutes the main theoretical contribution and attempts to combine two perspectives to create a new, usable theoretical tool. An existing theory of change framework (DuBow & Litzler, 2018) and existing guidelines on utilization-focused evaluation processes (Ramirez & Brodhead, 2013) are combined to enrich the field of program evaluation with a new framework that can be tailored to specific areas. Moreover, the study compliments the field of research in the nonprofit sector by providing a usable tool. The framework's contributions to the literature also hope to answer the call of Williams (2014), who initiated this research.

Furthermore, the study will contribute to the business world. Indeed, as it is known that evaluations in the nonprofit sector are not always conducted with a professional evaluator (Williams, 2014), relying on a framework is often needed. Thus, developing a helpful,

practical framework in cooperation with nonprofit organizations and for nonprofit organizations will hopefully contribute to advancing strategic evaluations in this sector.

#### **1.4 Thesis structure**

The thesis is constructed in five sections. The study starts with this introduction, highlighting the research gap and the reasons for choosing the study and the nonprofit field. In addition, the research question, some methodological information, and the dissertation's structure are presented. The second section depicts the study's theoretical background and concentrates on the necessary features, including program evaluation, strategic learning, theory of change, and utilization-focused evaluation. The third section outlines the methodological choices, data collection, and analysis methods. The fourth section assembles the results from the qualitative research. Finally, the findings are canvassed to synthesize the key discoveries and identify possible theoretical and practical implications. This chapter also discussed the study's limitations and suggestions for future research. This thesis considers the terms *evaluation* and *program evaluation* interchangeably.

## 2 Theoretical background

### 2.1 Understanding strategic learning processes

Strategic learning is part of strategy-making in strategic management and allows organizations to adapt (Starbuck et al., 2008). But before diving into strategic management, strategic planning processes, and strategic learning, it is essential to introduce the concepts and specifics of strategy. Quickly defined, strategy is about difference; companies aim to be competitive through their set of activities and their combination (Porter, 1996). Additionally, it is said that the ultimate mission of any strategy is to achieve the firm's mission within a designated time frame (Hastings, 1996). However, strategy is not a one-line road; there are different approaches, and many schools, visions, and explorations have been made of the concept of strategy by researchers and managers alike (Mintzberg & Lampel, 1999). Even further, and diving into the specifics, strategy is composed of three different branches, and this understanding of strategy is often referred to as the exploring strategy framework (Figure 1) (Whittington et al., 2023).



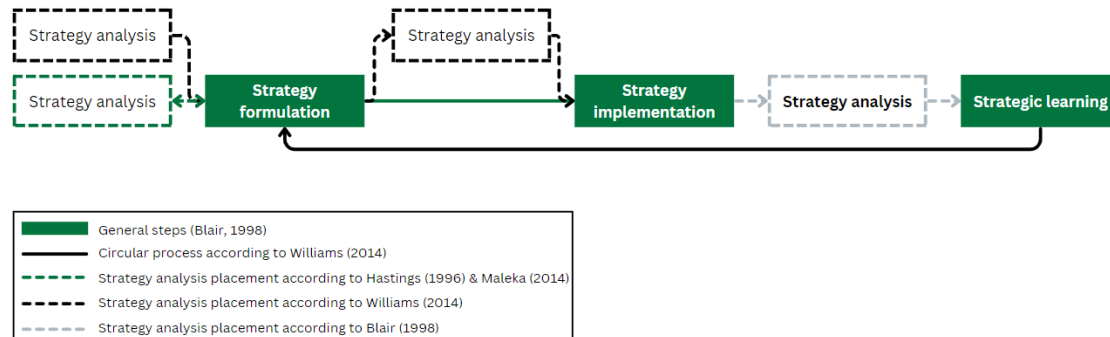
Figure 1. The exploring strategy framework (Whittington et al., 2023, p. 13)

As represented above, each strategy branch is responsible for different strategic roles. One branch aims to understand the strategic position and mainly deals with the internal and external context. Another one assesses the strategic choices for future decisions, such as the strategic options available for companies. Finally, one branch manages the strategy in action, mainly its implementation processes and evaluation. (Whittington et al., 2023). It is on the latter branch that this research will focus, and more specifically, on the strategy planning process and the strategy evaluation.

The strategy planning process holds both a symbolic and an instrumental role. Symbolically, it informs the different levels of the firms and represents a tool for communication; instrumentally, it helps companies to understand and analyze the continuous flow of information and data received (Dutton & Duncan, 1987) and provides a map to achieve strategic goals (Maleka, 2014). Many tools have been created throughout the years to support organizations and their strategic management development, but amidst all the existing ones, strategic planning is mainly used (Wolf & Floyd, 2017). Strategic planning guides companies into aligning visions and goals (Maleka, 2014), formulating problems and objectives, and choosing which strategy to implement, which grants it an essential place in the tools used by organizations (Wolf & Floyd, 2017). Over time, many frameworks have been developed detailing the steps and activities of the strategic planning process. Although it sometimes shows discrepancies in the variety of published models (Maleka, 2014), strategic planning is considered innovative and flexible for companies to use (Blair, 1998).

Figure 2 details the general steps that appear within the strategic planning process. Strategy formulation, strategy implementation, and strategic learning are the steps that create a consensus among authors (Blair, 1998), but two core differences are notable in the literature and can be seen representatively below. Indeed, it is noted that while some authors identify the process of strategy as linear, others add the notion of recurrence, making the connection between the final steps and the first one, creating a loop of never-ending re-evaluation of the strategy (Williams, 2014). The second discrepancy

relies on the placement of strategic analysis, which has sparked debate among authors, and presenting all three opinions is necessary.



**Figure 2. Strategic process and strategy analysis (Blair, 1998; Hastings, 1996; Maleka, 2014; Williams, 2014)**

Firstly, some authors note the importance of evaluating the strategy before the strategy formulation (Hastings, 1996; Maleka, 2014) because reviewing the quality of the strategy before its implementation would allow the firm to shift towards another strategy, a better-suited or better quality one, if the strategy analysis would result in outcomes judged negative, bad or not as good as another strategy's conclusions (Hastings, 1996). Secondly, it is justified that the strategy evaluation should be the final step in the process, thus allowing a better assessment of the resource allocation and the measurement of outcome (Blair, 1998). Finally, points were made about continuously evaluating strategy (Quinn, 1981) or at least at the pre- and post-formulation phases (Williams, 2014). Being able to evaluate through different phases of the process could reveal crucial information to decision-makers, such as how well the strategy would work or is working, the implications for the company or the resources needed for the process, or offer a better view of the possibilities, like those offered by emergent strategies (Quinn, 1981).

However, strategic planning processes can sometimes be misunderstood, misused, or wrongfully developed and implemented (Maleka, 2014). The strategic thinking phase can lack crucial connections, like the ones between assessing and adapting the strategy, not easily made by the organizations, but this is where evaluations for strategic learning

attempt to connect the broken links (Figure 3) (Williams, 2014). According to Hastings (1996), the quality of the strategy will lead either to the success or to the failure of the strategy in place, which evaluation can help with. It appears that to follow the strategic planning process, strategic learning is inevitable; it is considered to be the step that will close the loop of the strategy-making process (Williams, 2014) and similarly is the final phase of assigning the task of evaluating the results of the strategy (Blair, 1998). Such evaluation can focus on two aspects: the effectiveness of the whole strategic planning process or the usefulness and responsibility of programs (Blair, 1998). Strategic learning is needed to gather all the information and data from the previous strategic phases and turn the knowledge into lessons for decision-making (Coffman & Beer, 2011; Williams, 2014).



**Figure 3. Common broken link in strategic thinking (Williams, 2014)**

One helpful tool for organizations to know how qualitative their strategy is is to follow the evaluation process (Hastings, 1996), although the method is still considered a somewhat emergent discipline (Williams, 2014). Additionally, it is not the only tool that can lead to strategic learning; organizations can choose to follow another road, but conducting an evaluation will offer different results than other strategic learning methods could (Coffman & Beer, 2011). Indeed, such a method proved more important as evaluative thinking and reasoning reach more extensive conclusions than traditional reports (Patton, 2008a). Moreover, when deciding on evaluation as a strategic learning opportunity, organizations are ready because evaluation is not for all situations and requires internal staff, leaders, and situational, purpose, and resource readiness (Williams, 2014). Finally, evaluating a strategy integrates various partners into the process: the evaluator, of course, and the decision-makers and the staff (Coffman & Beer, 2011; Williams, 2014). Even though the leader will reflect on the questions the evaluation answers (Coffman &

Beer, 2011), it is necessary to incorporate the people into the evaluation process because evaluation is important at all levels. Everybody can judge and decide their actions (Patton & Campbell-Patton, 2021a).

## 2.2 Understanding program evaluation

As easily defined as possible, being evaluated or evaluating programs is about asking or answering questions (Patton & Campbell-Patton, 2021b). However, it is not as easily interpreted as it seems, knowing that the concept of evaluation integrates the notion of *value*. As individuals view it differently, program evaluation is subsequentially influenced by the particular group involved. The detailed definition of program evaluation may differ from one author to another. However, an agreement is reached by considering that it a tool used to judge if intended goals were met while also demonstrating the usefulness, worthiness, merit, and importance of the program to decision-makers through the collection of various data and information (Benjamin et al., 2023; Msila & Setlhako, 2013; Patton & Campbell-Patton, 2021a; Schwandt, 2015).

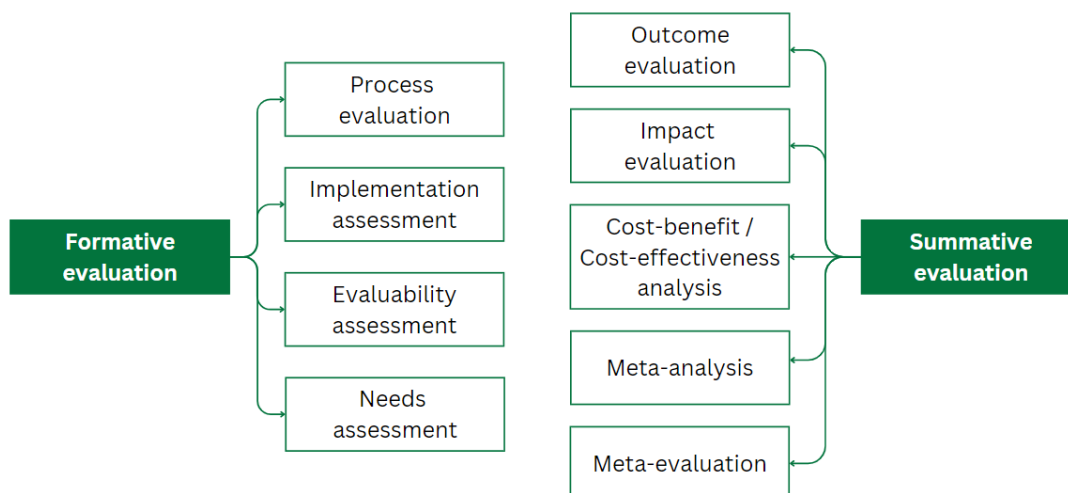
Furthermore, the interpretation of program evaluation can be discussed through two lenses. The pragmatic lens relates more to the definition of the factual term and uses evaluation as a literal means to an end; such a view helps to identify if and where funding should be directed (Patton, 2008a). On the other hand, the vision lens offers a broader perspective and conceptualizes evaluation as a piece that forms a new experimenting society (Patton, 2008a).

Evaluation and evaluating are not new concepts; the idea is essentially born from the meeting of knowledge and action (Patton, 2008a). The word itself, *evaluation*, refers to the old Latin language (Giancola, 2020a) and identifies the roots and the use of evaluation as dating back almost to the beginning of the world (Giancola, 2020b; Madaus & Stufflebeam, 2000). However, when thinking about program evaluation in particular, it is proper to admit that some aspects have appeared more recently or were given a clear

definition only in recent years; these are the profession of the evaluator, considered dynamic but still emerging and immature (Giancola, 2020b; Madaus & Stufflebeam, 2000) and the use made of the evaluation (Giancola, 2020b).

Evaluation, as a concept, appeared as a means to identify what is worth doing and where investments are worth placing, or not, only towards the late nineties following the realization that money cannot solve all problems and that, in the unlikely case there would be enough monetary support, human and social interactions also require more efforts (Patton, 2008a). Since its first boom, program evaluation has evolved and influenced generations; although the topic seldom appears in papers' headlines, it is an intense matter of concern for the people directly or indirectly involved in the program being evaluated (Rossi et al., 2017).

Program evaluation involves several approaches (Blair, 1998) falling under specific terminology (Giancola, 2020a); depending on the evaluation's intended goal, the evaluator takes different approaches. Figure 4 offers a schematic representation of the two different evaluation approaches, each linked to their terminological term, for a better understanding of what each term entails. Nevertheless, the view on the evaluation approach is not singular; it is noted that other authors view formative and summative evaluations differently, sometimes intertwined, than the visual representation offered here (Figure 4) (Chen, 1996). Adopting the vocabulary of most evaluators (McDavid et al., 2013a), evaluations will be considered as leading to formative or summative decisions (Giancola, 2014) to facilitate understanding for this dissertation. Formative evaluations, also called formative assessment by some, are considered by the evaluators if the evaluation aims to inform the program staff on potential improvement while the program is running, which is an important aspect here to consider; the other approach, summative evaluations, or summative assessments, are preferred if the aim is to discuss the program's future (Giancola, 2020a). Under each terminology, the evaluator can opt for several approaches falling under either of the aforementioned evaluative approaches.



**Figure 4. The program evaluation umbrella (Giancola, 2020a)**

Formative evaluations offer four approaches, in other words (i.e.) four ways to indicate improvement lines to working staff (McDavid et al., 2013a). Process evaluation regards the program processes and analyzes if the program works as intended, its delivery method, effectiveness, implementation, and the resources used (Blair, 1998); it can also be used mid-course to adjust the program design (Giancola, 2020a). Implementation assessment examines the program components, activities, and output to identify if the program was implemented as initially intended (McDavid et al., 2013a). Evaluability assessment can be considered a type of market research; the aim is to provide the evaluator with information, such as assessing the need or feasibility of the evaluation, that would help tailor practical, relevant, and feasible evaluation (Wholey et al., 2010). Finally, needs assessment evaluations are generally conducted before the start of a new program or before launching a restructured program; they are used to identify the needs of participants and stakeholders and determine how to meet both (Giancola, 2020a). Such evaluation aims to provide information on needs and recognize which program might answer those needs (EvalCommunity.com, 2024). Nevertheless, it is noted that the different kinds of formative evaluations might lack reliability if stakeholders are not enthusiastic about the evaluation or could cause constraints of time, resources, data analysis, or generalization of results (EvalCommunity.com, 2024).

Summative evaluations are conducted through five approaches. Outcome evaluation considers the results as proof of successful or failed objectives compared to the initial agreement; it is often considered more problematic as it requires identifying exactly what outcome resulted from that exact program (Blair, 1998). Impact evaluation also considers the program's results but from a broader perspective, including the impact of both intended and unintended goals in the evaluation (Giancola, 2020a). Cost-benefit and cost-effectiveness analysis particularly regard the monetary aspects of the program; they compare either the costs and benefits of the program in monetary units (cost-benefit) or the costs versus the outcomes (cost-effectiveness) (Giancola, 2020a; McDavid et al., 2013a). Both cost-related evaluations can be done before the program launch; they are then called *ex-ante* and compare the expected cost with expected benefits or outcomes, or after the program is implemented, they are then called *ex-post* and compare the actual cost with actual benefits or outcomes (McDavid et al., 2013a). Finally, a meta-analysis evaluation combines multiple studies' findings about a type of program to estimate the overall effect a similar program may have (Giancola, 2020a). It is distinguished from a meta-evaluation, which evaluates other, one or multiple, completed evaluations (McDavid et al., 2013b).

Program evaluations are generally conducted to provide data that would help decision-making (Benjamin et al., 2023). As such, the results are often considered the only part of an evaluation that can provide information, but more aspects of the evaluation have been identified as usable and informative, and these key resources are available for a wider audience (Weiss, 1998). Distinguishing clearly the use from the purpose, and both apart from the objective, of an evaluation is somewhat tricky because the terms are linked and often interchangeable in English (Schwandt, 2015). However, the literature identifies five utilizations of program evaluation: the symbolic use (Falconer-Stout & Jones, 2021), the instrumental use, the conceptual use, the persuasive use, and the influential use (Weiss, 1998); and various actors, intended and unintended users of evaluation (Patton & Campbell-Patton, 2021a), purposefully using the evaluation either productively or wrongfully (Patton, 2008a). Discussing the recognized uses and users of

program evaluation is especially important for this dissertation as the focus is on non-profit organizations that involve numerous users and require different types of conclusions that must be articulated around the primary use for the intended user (Benjamin et al., 2023).

### **2.2.1 Using program evaluation**

The literature recognizes two ways to use program evaluation, either for its content and outcomes (McDavid et al., 2013c) or as an entity, meaning that the sole idea of conducting an evaluation is usable (Weiss, 1998). Analyzing the utilizable content of evaluations, the most obvious answer in the literature to the question of “what is used in evaluation?” is the results (McDavid et al., 2013e). However, even for such an evident aspect, research proved that the results may be only partially used (Benjamin et al., 2023), wrongfully used (Patton, 2008a), or utilized without reaching a successful outcome (Williams, 2014). A study by McDavid and Huse (2012) provides answers to further details about what is concerned with the term results. The study presents the results as the evaluation's conclusions regarding program outcomes, for instance, efficiency and effectiveness, and the recommendations on the evaluated program processes, such as policy and budget decisions (McDavid et al., 2013c). Additionally, insights, ideas, and generalizations that often emanate from the results are also sections that can be utilized by the people related to the program or the policies that are evaluated (Weiss, 1998). Another way to use the evaluation concerns its structure, its design and its metrics; such methodology may help program directors to support their decision-making (Weiss, 1998).

However, as mentioned, the idea of conducting an evaluation can also be exploited positively and negatively (Patton, 2008a). Conducting a study is often considered an accountability tool to prove the worthiness of the program (Giancola, 2020a), sometimes described as symbolic use of evaluations (Falconer-Stout & Jones, 2021); it is shortly resumed one could say “the program is being evaluated, so it must be worthy”, or to influence the program and the policies in place (Weiss, 1998), or to decide on the program’s

future (Benjamin et al., 2023). Unfortunately, the executive may also use the fact that an evaluation is in progress as an excuse for inaction; the consequence of this mistreatment would lead the program directors to abuse the evaluation purpose to postpone the demand or the need for change (Weiss, 1998). This is often called the instrumental use of program evaluation (Benjamin et al., 2023; Weiss, 1998). Nevertheless, executing an evaluation may yield positive results for participating staff (DuBow & Litzler, 2018). As they begin thinking more about the goals to be accomplished and open internal conversation about the program's goals, clarification on what the program is and does appear while cohesion within the staff is created (Weiss, 1998). The reach is also likely to extend further and create incentives for improved performance to more than participatory levels (Wholey, 1993). This relates to the conceptual use of evaluation. Both the instrumental and conceptual uses of evaluation are directly concerned with the results and conclusions reached by the evaluation (Weiss, 1998).

Continuing the conversation, more dramatic or controversial uses, also called misuses, of the evaluation, are plausible, and two lines of conduct are opened by authors (McDavid et al., 2013c; Patton, 2008a; Weiss, 1998). The persuasive use means that evaluation results are manipulated to convey agreed-upon agenda and biases (Patton, 2008a) or to justify the program leaders' management type (Weiss, 1998). Sometimes, evaluations commissioned face defensive program managers, and the accuracy and completeness of information may suffer (McDavid et al., 2013d). Additionally, some evaluations are conducted while conclusions are already known or when executives have exiting knowledge of the possible outcomes originating from results, but support is lacking; thus, managers may choose to have an evaluation conducted to prove their already known points (Weiss, 1998) causing a distortion of results that supports pre-agreed decisions (Patton, 2008a). Such results tend to be used as a persuasive point to legitimize the executives' viewpoints and positions to try and gain adherents to their beliefs (Benjamin et al., 2023).

Meanwhile, the influential use of evaluation is expressed as a more manipulative utilization of evaluation and sees the conclusions being used to influence other parties, such

as institutions or events, or even sometimes work as a propaganda tool against public criticism (Madaus & Stufflebeam, 2000). The influential use of evaluation is also a known issue amongst evaluators. The persuasive and influential uses of evaluation are concerned not only with the results but also with the evaluation as an entity; the action of conducting the evaluation is a usable point (Weiss, 1998).

While the misuse of evaluations could seem far-fetched and does not represent what most evaluators expect as their work's ultimate goals, it is worth mentioning here because conclusions of evaluations are often separated from the organization conducting the evaluation and reach academics and other professionals who can use those conclusions to influence each other or the program evaluation community or impact policies, agendas and beliefs (Weiss, 1998).

### **2.2.2 Users of program evaluation**

Regardless of the chosen approach, the question of who uses the program evaluation is interesting, as understanding who the users are is key to preparing the evaluation (Patton, 2013b). Even more so, the evaluation customers are progressively becoming key factors in the evaluator's work rather than focusing on the decision-making level (Weiss, 1998). Overall, two groups can be identified. The people directly impacted by the evaluation (McDavid et al., 2013c; Weiss, 1998) and the others (Weiss, 1998). The expected primary users are the leaders, the decision-makers (McDavid et al., 2013c), the funders (Benjamin et al., 2023), and finally, all others inside the organization (McDavid et al., 2013c). However, it is also noted that the overall organization is often forgotten, although it is a crucial user to consider (Weiss, 1998). Because of its direct link to the overall goals, missions, strategy, and financials (Benjamin et al., 2023), and due to the use made of the evaluation regarding the policy-makers (Blair, 1998), and other limits imposed by law (Weiss, 1998), the overall organization is engaged directly with the evaluation. To an extent, the second group of users is considered with other similar programs, government or foundation representatives, policies makers, scientists, or even the public; the

evaluation does not directly impact these people but may have their own interests in the evaluation (Weiss, 1998).

Disregarding the evaluation approach taken and the users' goal, integrating the people involved in the program is highly recommended, as previous research demonstrated that involvement may lead to a better adoption of the evaluation (DuBow & Litzler, 2018). Now, it is essential to note that there are differences between theory-based and utilization-focused evaluations (Giancola, 2020a), the two lenses extensively discussed in this dissertation, and this aspect must be distinguished. *Theory-based evaluation* is an organized paradigm about what evaluation represents (Schwandt, 2015). Conducting a theory-based evaluation is argued to be more responsive to program needs and more relevant, thus ensuring utilization (Christie & Alkin, 2003); it does not consider the program as boxes to fill but asks questions to understand causality and relationships to understand *how* program functions, or not (McDavid et al., 2013b). It is also recognized that involving staff in a theory-based evaluation forces people to examine their assumptions about the program's processes (Weiss, 1998).

However, other authors contradict the unique use of theory-based evaluation. Indeed, counter-arguments exist again using one specific theory to evaluate one program, mainly due to questions revolving around the validation and methodology of using a singular outlook (Scriven, 1998; Stufflebeam, 2001). It is also discussed in the literature that the use of the results originating from the evaluation often falls flat as organizations welcome evaluations but do not synchronize them around strategy development, thus creating a gap between evaluation and strategy (Williams, 2014). Finally, a key author in program evaluation, Patton (1994), explains that organizations are co-evolving within their environment. However, traditional approaches to evaluation consider the program's structure as static, which allows the creation of logic models but may not make such models plausible or useful in co-evolving settings where objectives and environments are constantly changing (McDavid et al., 2013d). Thus, the utilization-focused evaluation approach will also be discussed. The reason behind choosing this particular

approach relies on its core focus, centered around the intended users and uses (Patton & Campbell-Patton, 2021a), and the role the evaluator has in working with internal stakeholders, which is discussed by McDavid et al. (2013d) as a stream of evaluative knowledge.

### **2.3 Theory of change**

The theory of change (ToC) falls under the spectrum of theory-based evaluation (Rogers, 2007), which appeared as part of the evaluation research field around the early 1990s (McDavid et al., 2013b) as a gap-filler in evaluations related to programs and policies (Blamey & Mackenzie, 2007). Before diving into the particularities of the ToC and its different representations, it is essential to discuss the specifics of theory-based evaluation because the topic is getting increasing attention in the literature, and with it, new terminology proliferates significantly (Rogers, 2007; Weiss, 1997). This paper will also try to provide a singular vocabulary to facilitate understanding of the concepts. Since its first mention in the late nineties, theory-based evaluations have become commonplace and are widely used in projects and companies of all sizes, but the terminology proposed for theory-based evaluation is fundamentally scattered (Rogers, 2007). Papers discussing the theoretical approach to evaluation will use numerous terms, sometimes interchangeably, in the publications (Rogers, 2007). Terms such as theory-based evaluations, but also theory-driven, theory-oriented, evaluation-theory, evaluation-practice, intervention theory, implementation theory, logic models, program theory, and theory of change (Brousselle & Champagne, 2010; Christie & Alkin, 2003; Coryn et al., 2011; Rogers, 2007; Weiss, 1997) and there are likely more terms that exist.

For clarity purposes, the reader is advised that this paper uses theory-based evaluation to discuss the general theoretical approach to evaluation, the theory of change for this specific approach within the general lens, and logic models as the chosen representative model. The terms are not used synonymously but as different concepts.

In literature, the ToC can be seen through various modeled representations, and terms are unclear as to whether specific models fall under the ToC view or a new theory altogether (Blamey & Mackenzie, 2007). Authors often refer to particular models as representing *a* theory of change (Patton, 2013b) or mention *theories* of change in the plural as they depict different models (Blamey & Mackenzie, 2007). The most common models existing as representations of a ToC are the logic models (Blamey & Mackenzie, 2007; DuBow & Litzler, 2018; Rogers, 2007) and the realistic evaluations, although this model is said to lack clarity (Blamey & Mackenzie, 2007). Despite this research focusing on the logic models as a ToC, introducing the other possibility is interesting for research purposes on program evaluation processes. Logic models are built to visually represent the program's theory and explain how the interventions should achieve the intended outcomes (Brousselle & Champagne, 2010); different logic models can approach different topics and theories (McDavid et al., 2013b). Alternately, realistic evaluations do not focus on the box-and-arrow explanation of programs' interplays but on the "psychological and motivational responses leading to behaviour change" (Blamey & Mackenzie, 2007, p. 446).

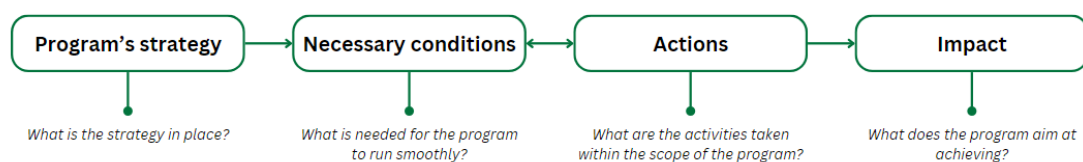
As a precursor to program and theory-based evaluation, Weiss (1997) mentioned that one reason such an evaluative approach may be advantageous is its possibility to explain better which outcome was the result of which process within the program. Although other emphasizes that the models used in theory-based evaluations often rest incomplete and fail to explain the causality relationships in the program components or the explanations of the outcomes reached (Rogers, 2007), it remains that the goal of theory-based evaluation is to assess the impacts and identify what works or not (Blamey & Mackenzie, 2007). As the ToC can have many forms and models, each evaluator may and should decide which theory of change is being implemented within the evaluation (Patton, 2013b). Additionally, evaluators in theory-based environments play a crucial role in the people involved and the program's context, such as its political, social, organizational, and individual environments (Blamey & Mackenzie, 2007). For these reasons, one model against another was chosen here.

The proponents of this theoretical approach to evaluation express that the theory would help organizations in three ways. Firstly, the ToC helps an organization prioritize, giving program staff a view of potential changes in order of importance (DuBow & Litzler, 2018). Secondly, the ToC assesses the expectations and feasibility of the program (Carman, 2010), i.e., it permits the comprehension of a program's model or the analysis of the relationships between the program activities and the intended outcomes (DuBow & Litzler, 2018). Thirdly, such an approach also generates conclusions on the program implementation to answer questions of accountability (Carman, 2010). Although the validity question remains when following theory-based evaluation, such a model would surely open the discussion surrounding the causal relationships between program practicalities (Weiss, 1997). Besides, following theory-based evaluation grants the contextual environment a central place; indeed, it is a key factor in understanding how programs and their effects interact (Blamey & Mackenzie, 2007), aiming at a deeper understanding of causality relationships in program components.

Nowadays, the ToC, as a tool of theory-based evaluation, is widely used amongst evaluators (Coryn et al., 2011; DuBow & Litzler, 2018), and when it comes to program theory specifically, the lead is most often granted to theory-based evaluations with a conclusion as a graphical representation of the components' relationships to each other (Coryn et al., 2011), representing one ToC model. When used adequately, this perspective on program evaluation carries several benefits (Carman, 2010). Through the ToC approach, stakeholders are guided to agree on the what and the why of their actions; the ToC helps outline the reasons for a program's initiatives (Weiss, 1995). Additionally, as the ToC confirms the allocation, or reallocation, of the resources on the program's key aspects (Weiss, 1995), the representation models vary greatly with more or less details and complexity (Coryn et al., 2011). However, the benefits of the ToC go beyond providing structure and accountability for evaluation. Indeed, an appropriately developed ToC model may help an organization in several other aspects, such as its decision-making purposes, direction, and priorities (DuBow & Litzler, 2018). External benefits may occur when evaluating the ToC as it leads organizations to agree on existing problems, which may help to influence

policies in place (Weiss, 1995). Nevertheless, one of the strongest assets of the ToC relies in it tailoring properties, as companies, with an evaluator, are able to develop their own ToC (DuBow & Litzler, 2018).

There are many ways to use the ToC (Breuer et al., 2014), and choosing one model against another means identifying several critical aspects of the program. The ToC approach is well-suited for more complicated programs, harder to evaluate or operating on several fronts because organizations developing this model will use it as a rescue tool that can inform on a program's planning and its development or help clarify vision and mission (DuBow & Litzler, 2018). Designing an evaluation following the ToC approach encompasses many steps, from planning to developing the instruments to conducting the evaluation, but literature is relatively scarce when it comes to identifying how the ToC is used to evaluate the program's implementation (Breuer et al., 2014; Coryn et al., 2011). Figure 5 presents a logic model (DuBow & Litzler, 2018) regarded as a great informative tool for external parties for program evaluation. Although models sometimes differ and do not always include all aspects, theory-based program evaluations usually highlight inputs and outputs, activities, initial and intermediate outcomes, and the impacts (Coryn et al., 2011). The logic model below, for instance, allows the program staff to pinpoint and understand better the link between what they do, i.e., the activities, and what they are supposed to achieve, i.e., the intended outcomes (DuBow & Litzler, 2018). For those reasons, this is the ToC representation chosen for the construction of the practical study, and it will serve as a base for building a more extensive framework later.



**Figure 5. Logic model as embedded theory of change. (DuBow & Litzler, 2018)**

However, although theory-based evaluation is an excellent tool for program evaluation and is argued to be more responsive and relevant (Christie & Alkin, 2003), it is also flawed in some ways (Carman, 2010). The frameworks are often simply intervention theory and not program theory, meaning they focus on activities rather than outcomes; the boxes and arrows schematic view created by organizations do not allow replications and provide little evaluative information (Rogers, 2007). Additionally, it is sometimes difficult for evaluators to measure multifaced contexts in which programs evolve (Blamey & MacKenzie, 2007). Furthermore, only using a ToC representation is insufficient when evaluating a program (Hollister, 2003), which is why this research will combine this model with another praised evaluation model, namely the utilization-focused evaluation theory.

## **2.4 Utilization-focused evaluation**

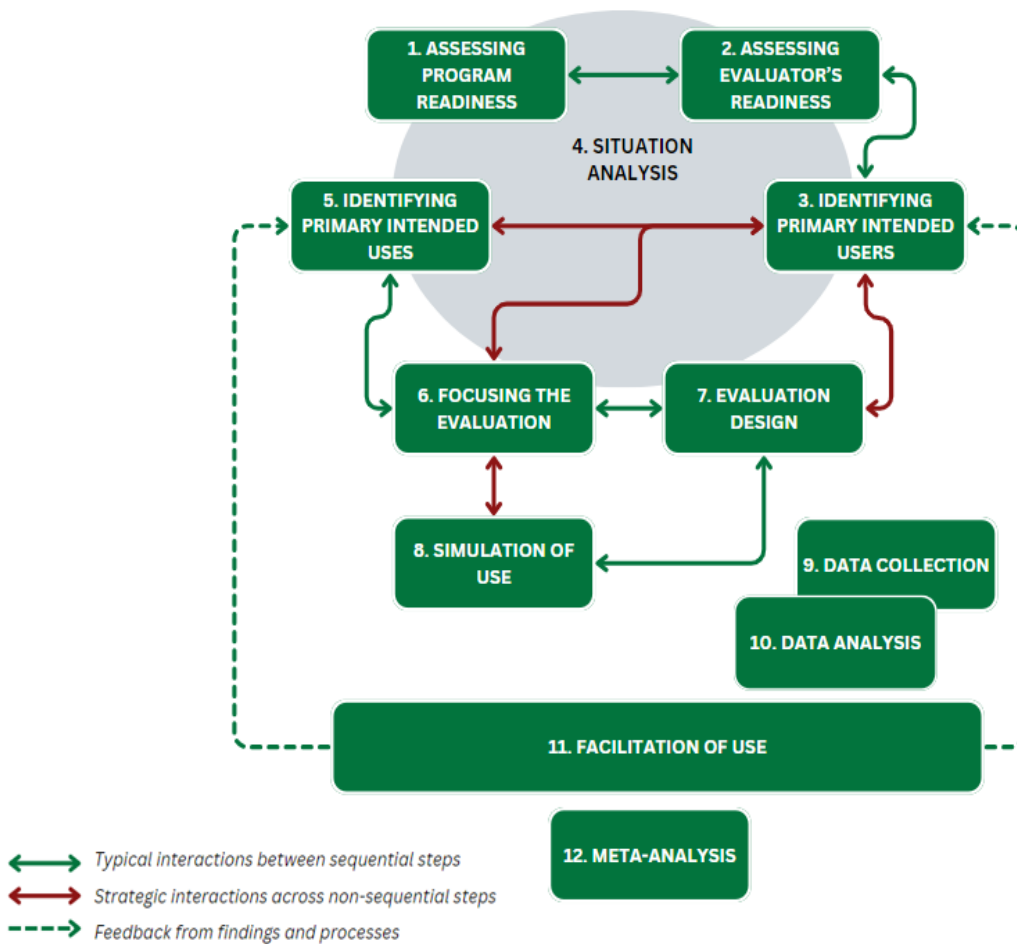
The Utilization-focused evaluation (U-FE) represents its lens to program evaluation, approach, and processes (Ramirez & Brodhead, 2013; Zaveri & Solomon, 2011) and does not fall under theory-based evaluation. This approach was developed by Michael Quinn Patton in 1978 and focused on evaluation that supports “intended use by intended users” (Patton & Campbell-Patton, 2023). This means the entire evaluation process is designed with the end in mind, considering the priority information and most important needs (Falconer-Stout & Jones, 2021). Because the U-FE theory is a process, it helps evaluators and other parties involved in the evaluation to select an appropriate methodology, tools, models, etc., to conduct the evaluation accordingly (Ramirez & Brodhead, 2013; Zaveri & Solomon, 2011). This is one of the most important aspects of the UF-E approach. To define the concept quickly, the U-FE aims to answer the questions of the quality and the use of evidence in decision-making (Patton, 2008a) while granting primary intended users and using a central role in the evaluation design and process (Patton, 2013b).

The general users and uses of program evaluation have already been discussed in the section on program evaluation. However, while the possible uses of evaluation remain similar for the U-FE as for general uses, it is important to dive further into the specifics

of the users as disparities exist. Indeed, the U-FE considers only “the primary intended users” (Franke et al., 2002, p. 14), representing a smaller part of the stakeholder group with a deeper commitment to the process and who are most likely to participate in the evaluation process. Unlike general direct users of program evaluation, said to be the leaders, decision-makers (McDavid et al., 2013c), funders (Benjamin et al., 2023), others inside the organization (McDavid et al., 2013c), and the organization as a whole (Weiss, 1998); for the U-FE direct users of the evaluations are the ones with decision-making power, and they are the individuals who hold the proper position in the evaluated environment to utilize the results (Franke et al., 2002). Michael Patton establishes the definition of the intended users of the U-FE approach through his publications and research and by others who have discussed or reviewed his work. The U-FE approach considers the values of the people responsible for applying the findings and implementing the recommended changes (Patton, 2013a). Franke et al. (2002) details further that the intended users can be a group of stakeholders but only if there are stakes and interest in the research and those who could make the most use of the results.

One of the core reasons why evaluators choose the U-FE approach is that it offers both freedom and a line of conduct. Indeed, the U-FE is a framework presented as a process to follow step-by-step (Figure 6), yet it provides evaluators with freedom thanks to the multiple ways of use possible (Wiréhn & Haag, 2020) and allows users to become co-designers of the evaluations (Ramirez et al., 2017). As Patton (2013b) well details, instead of being another strict methodology, the U-FE is a guiding framework with open possibilities and choices where the evaluator can include various evaluation methods “within an overall participatory paradigm” (Ramirez et al., 2017). Additionally, the U-FE processes to evaluate programs answer prior remarks in which authors have introduced the idea that using results does not often meet the required change and adaptation (Williams, 2014). The utilization of conducted evaluation results is said to be evident when four factors meet: 1. the evaluation design is fit for the case, 2. the intended users are involved and committed to the evaluation process, 3. the results presented are considered relevant by the users, and 4. the external environment does not conflict with the

given results (Franke et al., 2002). Thus, choosing the U-FE approach is a smart move toward seeing results used by those who can adapt and change based on recommendations.



**Figure 6. Using the utilization-focused evaluation processes (Ramirez & Brodhead, 2013)**

Utilizing the U-FE theory for evaluation is a process (Zaveri & Solomon, 2011) that its creator elaborates clearly through the publications and texts published. The U-FE framework can be found in various representations; initially, being a six-step process, it evolved to a twelve-step framework and can now be found with up to seventeen steps. Figure 6 represents the 12-step view as the chosen visual representation for this research. Through those steps, the evaluator is guided in the evaluation preparation, design,

conduction, and the following steps that could be needed. By using the U-FE approach to evaluating a program, it is important to remember to work with the primary intended users throughout the process (Patton & Campbell-Patton, 2021a; Zaveri & Solomon, 2011); evaluator and intended users are co-designers of any U-FE (Ramirez et al., 2017). The close cooperation provides the evaluator with insights into identifying the meaningful questions the users need answers to (Patton & Campbell-Patton, 2021a). The implications of the end use of the evaluation are part of every decision taken regarding the evaluation design, process, and completion (Patton, 2008b); it is regularly described in the process that users and uses should be reviewed, rechecked, and reconfirmed (Patton, 2013b).

The assumption of the U-FE theory relies on the real use of the findings by the primary and intended users (Franke et al., 2002). Having the intended users committed to the process is a core focus of the U-FE theory (Ramirez & Brodhead, 2013) because they must be included and participate in the evaluation process (Zaveri & Solomon, 2011). Additionally, this approach lets the evaluators grant intended users a central role. Initially, evaluators give intended users ownership over the evaluation process and findings, creating active participation, involvement, and use of the evaluation (Patton, 2008b; Ramirez et al., 2017). Towards the end, the approach allows evaluators to facilitate decision-making amongst the primary intended users (Wiréhn & Haag, 2020).

It is said that identifying the concerned parties would lead to a higher probability of having the results indeed used by those who need or request the evaluation results (Patton, 1997); the intended users of the evaluation then become the base of any evaluation conducted under the U-FE theory approach (Franke et al., 2002). However, an important point to be noted is that the focus is solely on users who are concerned by the results and the evaluation process, which does not mean all the users (Franke et al., 2002); knowing the who is crucial and may lead the evaluator into using a standard evaluation method or take the road to a more innovative means to evaluate; the core focus must remain however on who will use the generated findings (Patton & Campbell-Patton,

2021a). Once again, it is important to note that the U-FE theory is not an evaluation method by itself but rather an approach to inform the evaluator and the concerned parties on how to conduct the evaluation based on the intended users and on how the evaluation design and results are intended to be used (Patton & Campbell-Patton, 2021a). Evaluations must always provide relevant and understandable results, as organizations will care about what is recognized to impact them directly and what is within their power to control (Patton, 1997; Thompson, 1967). Following a U-FE provides just that (Patton, 2013a).

However, the U-FE approach can also be misused and has disadvantages. Because the U-FE considers intended users and uses in the entire evaluation process (Wiréhn & Haag, 2020) for situations where both are unknown, the U-FE is unfitting. Additionally, an oversimplification of the conclusions is possible; once the initial picture of the results is shown, decisions are taken mechanically without further investigations (Patton, 2008a). Imperfect results can also cause misuse of findings, trouble moving from findings to action, or ignorance of the findings altogether (Patton, 2008a). A less dramatic disadvantage is that the U-FE approach may tempt or trap the evaluators in some ways, impacting the work done with the evaluation; for instance, if the evaluator regards itself as a decision-maker, is coopted by others, targets the wrong users, or uses, or if assumptions on roles and power are made (Zaveri & Solomon, 2011). Finally, the evaluator's perspective versus the intended users' perspective may be questionable as both are valuable. Separation may occur, although this can also be considered a strength for tailoring the evaluation to make a bigger impact (Wiréhn & Haag, 2020).

## **2.5 Specifics of public and volunteer organizations**

As the research focuses solely on the public and volunteer community, a short description of their specificities is needed. Unlike for-profit organizations, nonprofits operate with different challenges, structures (Benjamin et al., 2023), pressures, and stakeholders (Carman, 2010). Even further, differences exist also between public and volunteer

organizations. One difference resides in their respective organizational form. While public organizations have consistently been recognized as a clear type of organization, it is only later that volunteer organizations were distinctively recognized as having an organizational form (Benjamin et al., 2023). Variable in form, with its unique challenges and effectiveness, the organizational structure taken by such organizations can vary in form, and the impact it has on engagement and authority, in return, affects both the values and the norms of the organizations (Benjamin et al., 2023).

Another difference links to strategic choices. Regarding strategy making, further discrepancies exist between traditional profitable organizations and nonprofits; the difference goes deeper and varies between the different types of nonprofits, public and voluntary, or in the private sectors (Maleka, 2014). Factually, nonprofits do not have the same available resources as private for-profit enterprises do (Rudov et al., 2017), but in all three areas, strategy-making differs due to the different pillars on which they each rely or focus (Maleka, 2014). In the private sector, strategy rests on competition and satisfying customers better than the other firms within the same environment (Porter, 1996). The public sector focuses on delivering goals and fulfilling the people (Maleka, 2014). The voluntary sector is the in-between; similarly to the private sector, it is subject to competition, especially for funds and donations (Benjamin et al., 2023; Maleka, 2014), and like in the public sector, the strategic decisions must answer set objectives (Maleka, 2014). However, regardless of the strategy chosen, similar strategic tools are used by nonprofits and for-profits, with arrangements and modifications.

Indeed, many changes have led public and nonprofit organizations to use strategic planning tools (Bryson, 1988). In recent years, and with the increasing demand to account for nonprofits' actions and results (Carman, 2010), whether it be new challenges that emerge from the continuously evolving market dynamics, the need for more services with fewer resources (Maleka, 2014), or the uncertainty of revenues and sources of income, the use of strategic planning has been increasing in both public and volunteers organizations (Bryson, 1988). In public organizations specifically, not only is it necessary

to complete the strategic planning steps and identify if economic development is well justified, but it is also important to have program evaluations to close the loop (Blair, 1998). Furthermore, and valid for all nonprofit organizations, there is an increasing demand for project proposals to include logic models or program theory to understand better the value of management, planning, and evaluation (Rogers, 2007). Unfortunately, justifying programs' implementation, actions, and successes is neglected and forgotten by many, especially in public organizations (Blair, 1998).

Another difference regards the tools to use. Remembering that the goal of evaluation is to provide information to decision-makers on current and future programs (Patton & Campbell-Patton, 2021a), for nonprofit organizations, the tool in use will depend on the strategy in place (Benjamin et al., 2023). For public and volunteer organizations, evaluation is a logical tool and step to ensure that efforts and resources placed in programs are helpful (Blair, 1998). Indeed, in nonprofit organizations, unlike for-profit organizations, the financial performance is not considered as an outcome but rather as an input, and they face multiple entities when assessing their effectiveness with sometimes different criteria used to determine (Benjamin et al., 2023). Thus, strategically choosing the tool according to the strategy is crucial as the results differ, and recommendations can vary on aspects such as staff allocation, cutting or renewing grants, or modifying the program's design (Benjamin et al., 2023). Even when choosing a specific tool, like program evaluation, nonprofit organizations must ensure to fit the settings, for instance, formative or summative evaluations, to the needs of their program (EvalCommunity.com, 2024)

## **2.6 Proposed conceptual framework**

Research on theory-based evaluations expressed that the existing models used to analyze programs through this lens remain incomplete, fail to explain the links in the program components, or do not represent how the outcomes are reached (Rogers, 2007). Additionally, the knowledge gained from publications on both the U-FE approach and the ToC demonstrated that the users are active in the evaluation process (Blamey &

Mackenzie, 2007). Thus, the conceptual framework (Figure 7) will try to fill the gap identified in theory-based evaluation research and complete the logic model representation with more information to detail as clearly as possible the relationships between the program components interacting in the nonprofit sector, and the reach of the outcomes while placing the users as a key part of the evaluation process. The conceptual framework summarizes the concepts discussed in the theoretical background and guides the research's next step to gain a deep understanding of program evaluation processes for strategic learning in the nonprofit sector.

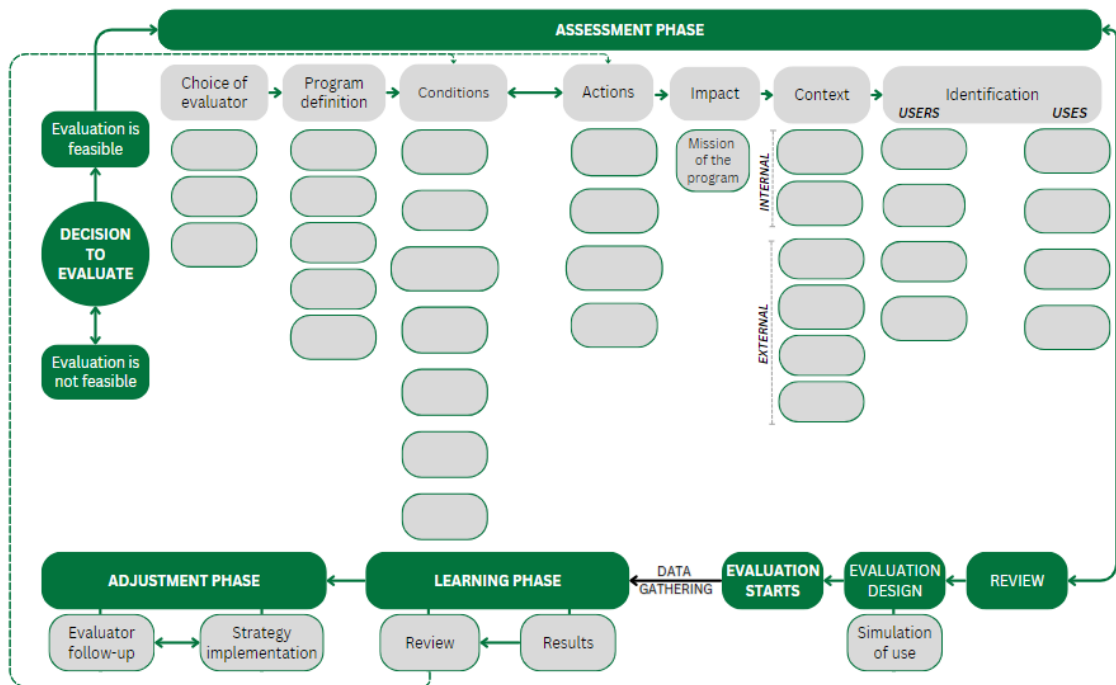


Figure 7. Initial framework

### 3 Methodology

#### 3.1 Research approach

The study approach was qualitative research that focused on public and volunteer organizations. The selection of the four cases for this study provided controlling variance that gave the study better data to answer the complex social processes linked to the thesis' research question (Eisenhardt & Graebner, 2007). Indeed, the nonprofits operate in a similar environment; they provide information and guidance to internationals living in Finland. Yet, they represent two different types of nonprofits: public and volunteer organizations. When conducting a study, it is crucial to select an adequate methodology because it is the fit that will guide the quality of the research (Edmondson & McManus, 2007). A representation of the methodological approach for this study can be seen in the research onion (Figure 8). The research onion was first presented by Saunders et al. (2023) and remained consistent with minor changes throughout the years. The model depicts the vital methodological process and decisions any dissertation writer should follow. Moving from the outer layer to the inner layer, the first step is identifying the research philosophies and selecting theory development approaches, methodologies, strategies, and time horizons. The study's design ends with the techniques and procedures for data collection and analysis (Saunders et al., 2023).

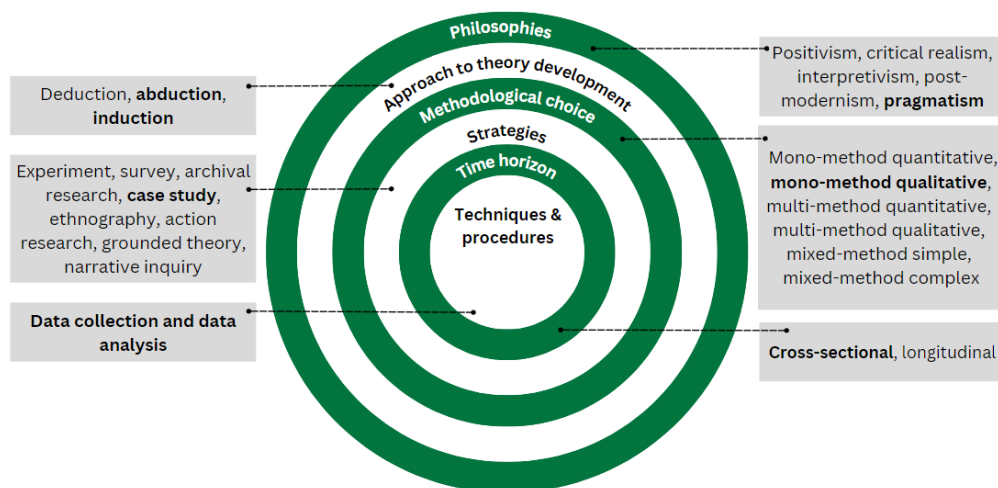


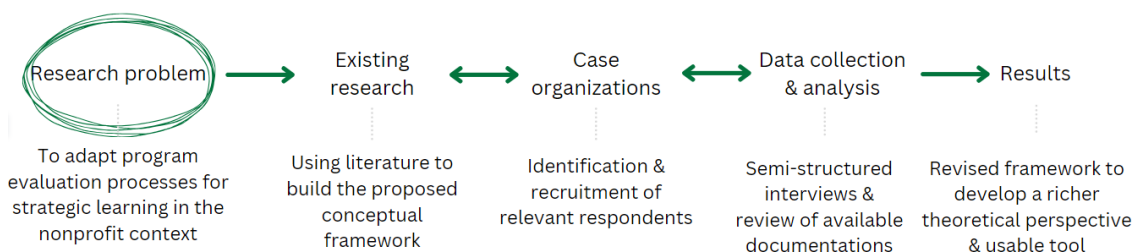
Figure 8. Research onion for the study (Saunders et al., 2023)

Starting with the first methodological decision, the definition of fundamental philosophies, it presents five ways one sees the world around, five options with their view of acceptable and desirable knowledge. Positivism and interpretivism are two mainstream views classified as historical, while pragmatist, critical realist, and post-modernist philosophical views on research have appeared more recently. This study adopts a pragmatist view, knowing that the goal is to make a difference in organizational practices. Additionally, this view starts with a problem and gives the research question a central role, which also drove this research in the first place. With the help of an adequate methodology, it considers the data collected, theories, hypotheses, and research findings not as abstract values but as actors of thoughts and actions in a specific context with practical implications. (Saunders et al., 2023)

Peeling the second layer of the onion, three approaches to theory development are presented. The deduction approach starts with the theory, and the study tests said theory. The abduction approach moves back and forth between data and theory to create or add to a theory and test its plausibility. The induction approach gathers data to build or add to a theory. (Saunders et al., 2023). This study's approach to theory development can be justified in two ways. Because the proposed initial conceptual framework was built using existing research as the base links to theoretical findings, and because regular movement between data and theory was clear and added to the two theoretical perspectives discussed in this thesis, the abduction approach is justified; however, the study does not test the theory's plausibility.

Additionally, because the study gathered new data to create a new conceptual framework adding to the theory of change and utilization-focused evaluation approach, the induction approach is justified; however, no new theory has been developed. As a result, it is possible to say that the approach to theory development for the study is a mix of induction and abduction, as existing research was used, but modifications and additions were made based on the data analysis outcomes. The qualitative data for this study generate unique insight into tackling the complex social processes linked to the research

question that quantitative data would not answer (Eisenhardt & Graebner, 2007). Figure 9 offers a view of the induction approach adapted to this study, starting with the problem answered in a revised conceptual framework after a thorough data collection and analysis process with suitable organizations. The research aims to add a richer theoretical perspective to the UF-E theory and the theory of change.



**Figure 9. Approach to theory development for the study (Saunders et al., 2023)**

Finding the methodological choice to implement for this research followed the recommendations given by authors in the field based on the previous decisions outlined in the research onion (Figure 8). As an inductive approach to theory development is decided, the logic presented for this dissertation is that the research drives the theory (Saunders et al., 2023); thus, a qualitative study was initiated. Furthermore, knowing that the research was conducted using a single technique of data collection, semi-structured interviews, the methodological choice is mono-method qualitative research. Indeed, the research will produce non-numerical data (qualitative) using one data collection procedure (mono-method).

The final inner layers of the research onion close the research design with the choice of strategies and time horizon. The research strategy links the philosophy choices with the methodology (Saunders et al., 2023); the case study approach was evident as semi-structured interviews were conducted with organizations. Furthermore, a multiple-case study provides more robust theoretical insights thanks to the varied empirical evidence grounding the theory building (Eisenhardt & Graebner, 2007); it is a multiple-case study that was implemented. The number of organizations to interview intended to respect the guidelines provided by authors in the field. Saunders et al. (2023) justify that with

inductive reasoning, a smaller sampling if better suited, while Eisenhardt (2007) remembers that each added case can multiply the reasoning power; thus, smaller sampling is typical in multiple-case studies. For these reasons, the study was conducted with the help of four organizations. Finally, as this study focuses on a given phenomenon in a given period, a cross-sectional study seemed the most suitable time horizon option (Saunders et al., 2023).

### **3.2 The case organizations**

A total of four organizations were studied for this research. The nonprofits studied for this research were either public or volunteer organizations, all based in Finland and working in the same environment, dealing with internationals in Finland. Validating two different types of nonprofits interacting in similar environments provided controlling variance for the research. More specifically, two public organizations and two volunteer organizations participated in the research, which allowed equality and hopefully led to better parity in answers given. In addition, receiving the view of two different types of nonprofits suggests enriching literature with broader views within this context of nonprofits. The reader may note that to provide anonymity to the respondents, the participating organizations are not presented in the order they were studied, and the interviewees were attributed a random number (appendix 2) not corresponding to any particular order; indeed, the order of the case studies is irrelevant to the successful reach of the study.

A total of two public organizations were participants in this study: MONI-Info Seinäjoki (Moni) and International House Oulu (IHO). Moni is a service part of the city of Seinäjoki that was created in 2020 and was initially planned to be a short-term three-year project running from 2020 to 2022; but due to its need and success, the project was established as a permanent operation in the city's affairs as in January 2023 (Seinäjoen kaupunki, 2023b). Additionally, internationality and international activities were given a clear role in the city's newly released strategic development report for the rest of the 2020s (Seinäjoen kaupunki, 2022). Ultimately, the city wishes to utilize and promote

internationalism to strengthen the city's attractiveness and competitiveness factors (Seinäjoen kaupunki, 2023a). The Moni service is designed to support internationality and immigration in two ways. Firstly, the service provides information and guidance to immigrants in the Seinäjoki region. Secondly, the service helps the city's actors with their questions on immigration. Overall, the service aims at offering high-quality service to all city residents without discrimination. At the time of writing, the program's services are offered free of charge to customers, and the program runs thanks to three employees employed by the city. All three employees have been interviewed for the research. As it is part of the city of Seinäjoki, Moni relies on a budget given by the decision-makers acting within the city's governing body.

The second public organization participating in this study is the International House Oulu (IHO). IHO is an advisory service created in April 2023 and based in the city of Oulu in the key location of Business Asema, where internationals and locals can find a wide range of answers, with up to five different informational services in one place. The IHO service has three full-time employees and other staff who partly work for the service. The IHO oversees several groups that oversee certain sub-services, making IHO the top of the umbrella under which several sub-services are found. Although independent with their own name and brands, services such as the multicultural center or guidance for business establishment, to cite a few, are branches found under the IHO umbrella. Overall, the IHO service informs several target groups regarding different information they might need; professionals and individuals can interact with the IHO team and staff. The IHO staff provides residents, international newcomers, and immigrants with daily life information on the city of Oulu, with guidance on requested matters, networking events, personal support, and friends-making.

Additionally, the IHO grants companies their knowledge regarding hiring international workers. Finally, the IHO organizes events and makes the city of Oulu a more thriving place for internationals living in Finland or Oulu. (International House Oulu, 2023). If the IHO cannot answer the customer, guidance on where to find the answer is always given.

The IHO staff proudly says that “the customer is always at the right place” when visiting their premises.

Furthermore, two volunteer organizations were participants in this study: Bloom Finland and the Multicultural Association of Porvoo (MAPorvoo). Bloom Finland aims to build a community where everybody feels welcome and internationals and locals can thrive together (Bloom Ry, 2023). The organization functions thanks to all its volunteers allocated with different roles and responsibilities, as well as the paying members of the association. At the time of writing, the association is composed of an executive board of eleven people who handle the general activity, ensuring the smooth running of the organization. Three local heads of operations accompany the activities of the associations to a local level, and each person responsible oversees one or several local volunteers to carry out the association’s events and activities. The nonprofit organization was created in 2020 by three internationals living in Finland and provides cultural and social events for its community and members. The Bloom Ry nonprofit association started in Tampere, but as the need for such services grew in the country, the organization has recently, in 2023, grown to open other branches, notably in Oulu and Porvoo.

The fourth participant in the study is MAPorvoo. The nonprofit organization is based in Porvoo and aims to build a multicultural community within their area, i.e., Porvoo and the southern Uusimaa region (MAPorvoo, 2017). The association was created in 2013 by immigrants integrated in Finland who noticed the lack of a community in which they could share their knowledge about integration between internationals and Finns. Today, the organization is still thriving and following the same structure as it was first created. To run smoothly, the organization relies on and functions thanks to three key actors. First, its governing body of volunteers who, at the time of writing, numbered four people. Second, it pays members and event attendees, the target group, to join the program they choose to run. Thirdly, the partners with whom they work will be able to provide services, events, and tools to facilitate the integration and relations between Finnish people and foreigners (MAPorvoo, 2017). To successfully achieve its mission and goals, the

organization provides regular programs, events, and projects to its members and guests who attend the activities.

### **3.3 Data collection**

Collecting data for this study was two-phased; combining data collection methods is particularly useful in case studies (Eisenhardt, 1989). The primary data collected for this study is the empirical results of the interviews conducted with the public and volunteer organizations. This data was collected through semi-structured interviews with a protocol of 18 questions and eight respondents (appendix 1). Such an approach to the interview allowed the interviewer to bounce on interesting topics and ideas raised by the respondents during the interview (Saunders et al., 2019). Each of the organizations was asked the same set of basic questions, but as the case studies all followed a semi-structured approach for the interviews, depending on the interviewee's answers, the order of the questions may have changed, and the information from previous interviews were considered when conducting the following ones. However, the main topics were discussed in all interviews. The questionnaire was tested to ensure that all questions were understandable for the interviewees, theoretically relevant, and could provide the most data. The questionnaire in the appendix is the final version; two other versions existed before the sample people approved the final one. The questionnaire was tailored with both open-ended and closed questions, and While the closed questions aimed at getting direct answers for building the revised conceptual framework, the open-ended questions permitted a more comprehensive approach to information gathering on the topic discussed (Eriksson & Kovalainen, 2008).

The interviews were conducted randomly between the public and volunteer organizations and with the maximum number of people possible to ensure relevant information was gathered. Each conversation was recorded, transcribed, and analyzed similarly and promptly after the data collection phase to ensure the proper collection of information through the interview; this also reinforced the validity of the research (Eriksson &

Kovalainen, 2008). The interviews were all conducted in English, as this is not the mother tongue of several respondents; it is possible that feelings and emphasis were not fully captured, or the interviewer or the interviewees misinterpreted the tone of the voice. The number of respondents varied from one to three interviewees per organization. Such number per case is justified as they represent the number of decision-makers, or in one case primary decision-makers, that each of the case organizations, both non-profits and volunteers, had within their presiding board or organizations. The cases were selected for their interest in the research and study, their similarity in terms of environment, i.e., they all deal with internationality in Finland, and for a shared parity between public and volunteer organizations.

Apart from the primary data, secondary data was collected through existing sources. The use of various data sources is supported by Yin (2009), who mentions that varying the sources would provide stronger reliability in a case study. Two kinds of secondary data were utilized for this research. Public documentations were encountered, such as published reports, public documents, strategy reports, websites, and other business newspapers found online. Additionally, some of the case organizations granted access to private documentation that supported the goal of this research. The private documentation shared included reports and findings of internal evaluation, access to the organization's database, and internal evaluation forms used by the organization and the staff. This dataset is not open to all and is only accessible internally. The chance to access such data allowed a deeper view of some aspects of evaluation. With the promise of anonymity and agreement from the concerned parties, some information has been used for this research.

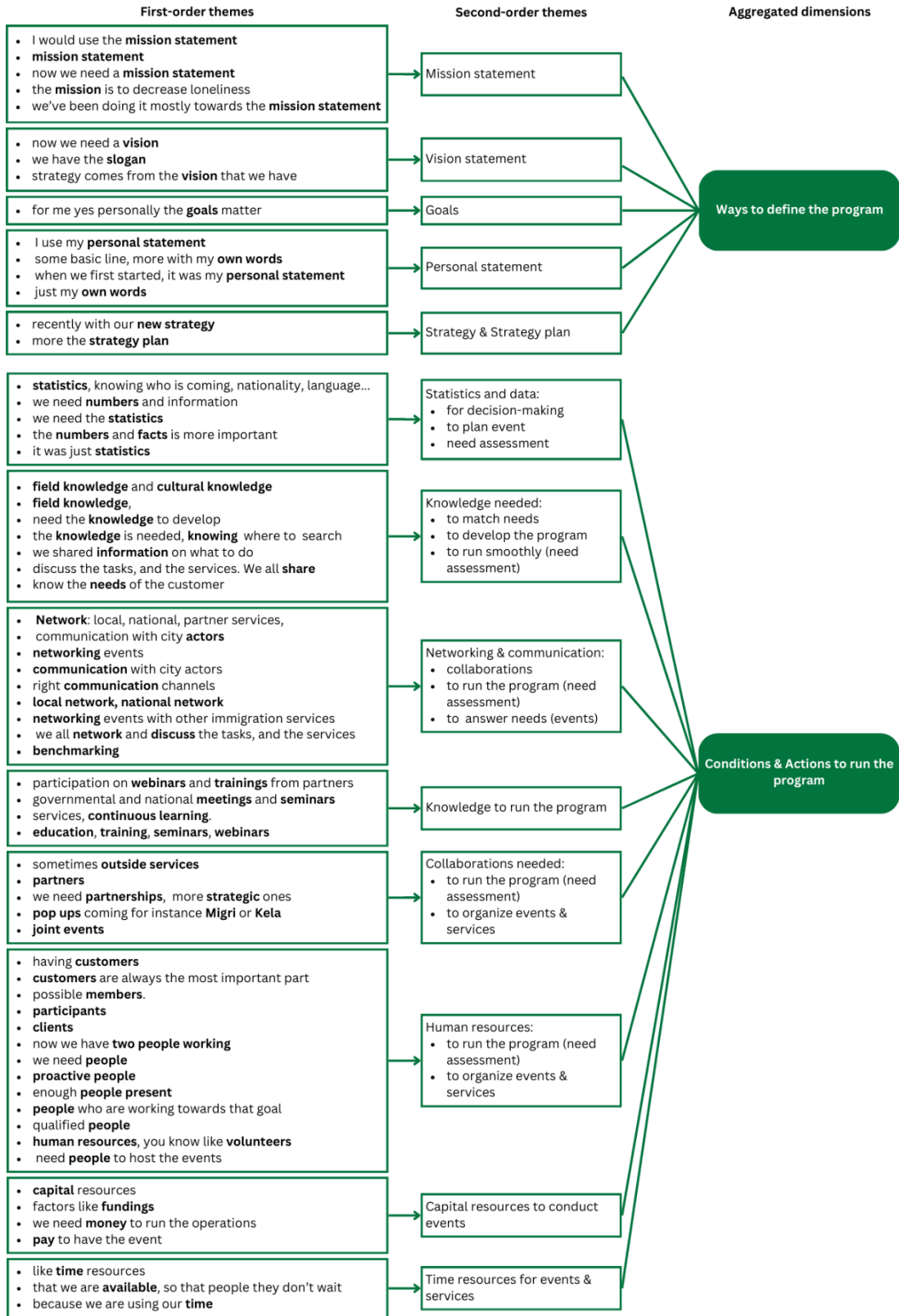
### **3.4 Data analysis**

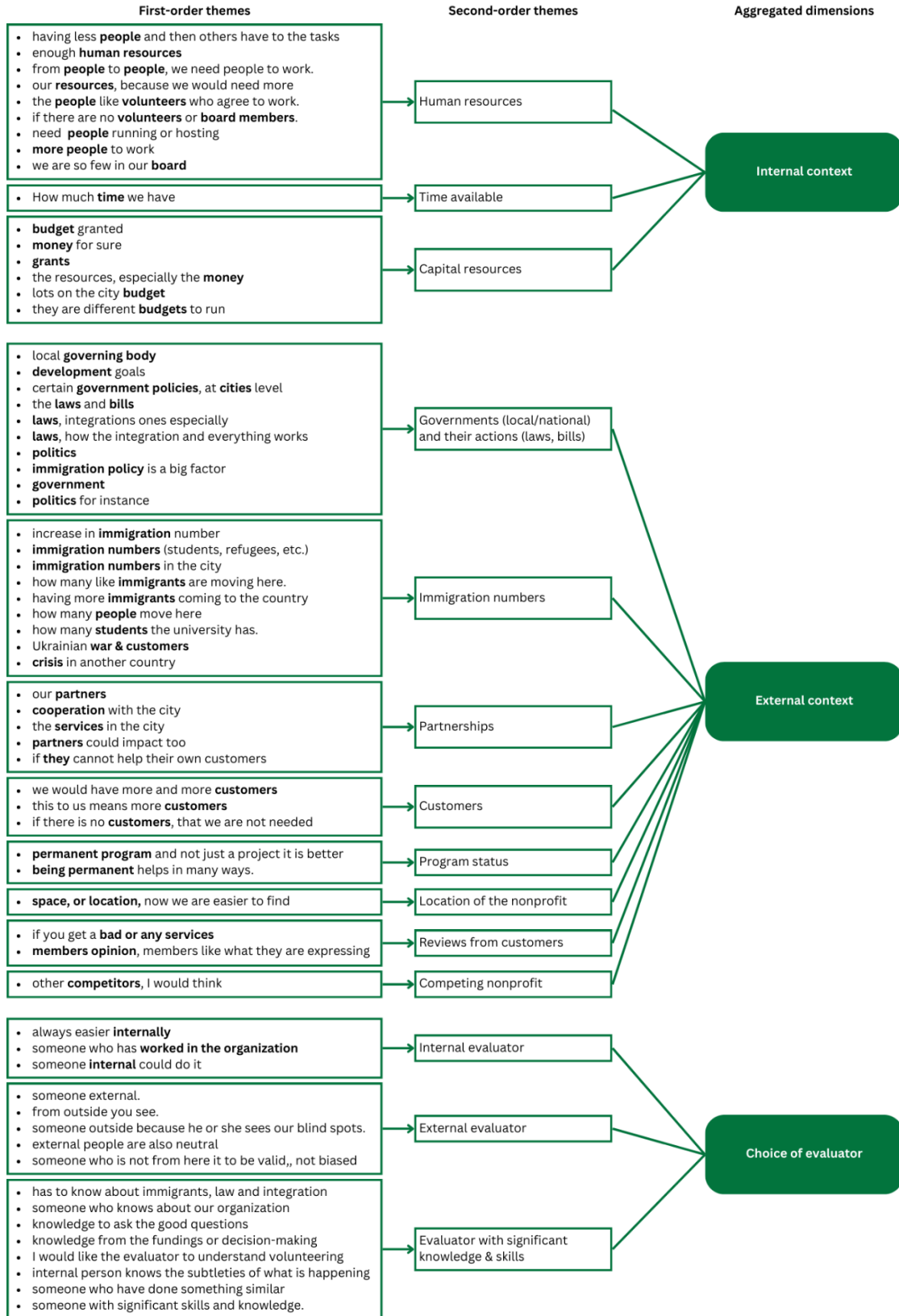
The empirical data aims at completing the initial framework presented earlier in this thesis. The framework was based on qualitative information and omitted quantitative data; this is justified because, although it is often numerical data that is used to evaluate

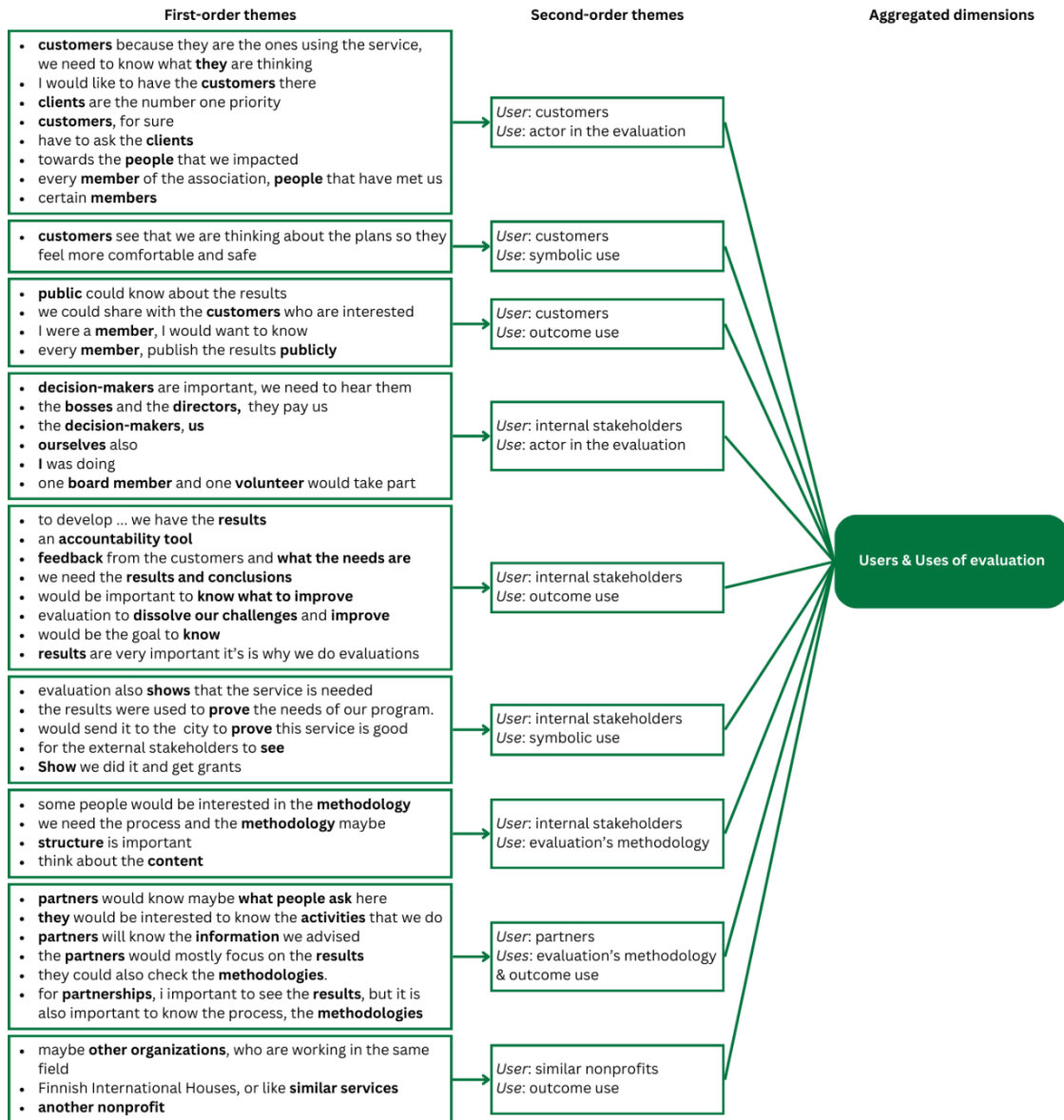
strategy, financial analysis may narrow the conclusions solely to financial returns, closing the conversation on the possibility of seeing a successful mission statement achieved (Hastings, 1996).

The study followed a mixed approach to theory development, which was respected for data collection and analysis. As both phases are intertwined, this led to analyzing several factors gathered while collecting the data via the interviews. Such factors are spoken communication, meaning the content of the interview, and unspoken communication, such as the interviewee's body language, possible emphases, or tone of voice observed while discussing the primary questions. The interview content was gathered and studied similarly according to a plan; following a clear plan for the data analysis is recommended to avoid facing a large amount of data that could become incoherent (Farquhar, 2012; Yin, 2009). Furthermore, as methodological experts recommend executing data analysis rapidly after the primary data collection phase (Eriksson & Kovalainen, 2008; Farquhar, 2012), the transcription of the interview content was adapted to the text format immediately after each interview. Additionally, any observation notes on unspoken communication taken during the interviews were combined with the transcribed text and placed where the interviewer noted it. The transcription file recorded a total of 50 pages.

It is critical to organize complex data gathered to classify and determine which themes were recurring in the data collection step (Saunders et al., 2019). Thus, analyzing the data using the Gioia methodology was justified. To accomplish this, the recurrence of common and frequent terms, phrases, words, and terminology used by the respondents were built into first-order themes. The second-order themes examined the dimensions of the first-order themes closer to identifying connections, links, or patterns. This step involves stronger complexity and abstraction due to the influence the theoretical interpretations and respondents' wordage have on the researcher's perspective (Gioia et al., 2013). Finally, aggregated dimensions were identified based on the themes and systematically compared to the initial theoretical framework (Gioia et al., 2013). Figure 10 presents the full data analysis structure.



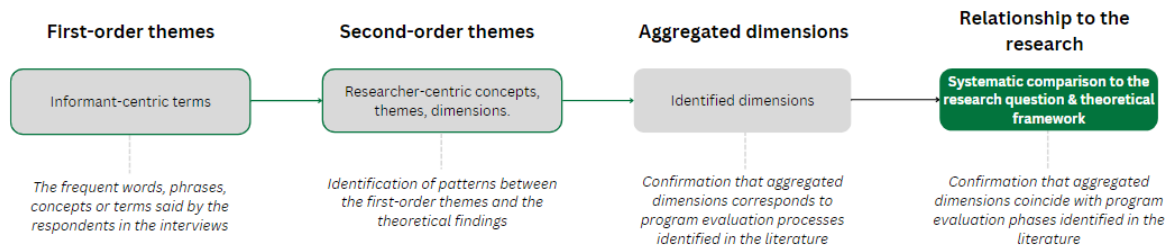




**Figure 10. Full data structure of the study based on Gioia's methodology (Gioia et al., 2013)**

As the research mixes two approaches to theory development, using existing research with new data, the findings are reported differently than the traditional representation of Gioia dimensions. This means that the aggregated dimensions identified in the study from the analysis of first- and second-order themes (Figure 10) are not new additions to theory, instead, they needed to coincide with theoretical findings. To clarify how the analysis led to the written findings and the final framework, Figure 11 is helpful. The aggregated dimensions correspond to the theoretical findings on program evaluation

processes, while the second-order themes outlined new concepts linked to them and are reported as such under the program evaluation processes, i.e., the aggregated dimensions. The written findings discuss recognized dimensions and their associated second-order themes. The final framework uses identical color-coding as the ones represented in Figure 11.



**Figure 11. Gioia's methodology model adapted to the study (Gioia et al., 2013)**

## 4 Findings

### 4.1 View and experience on evaluation

By the beginning of the interview process, the respondents had some understanding of what an evaluation was and basic knowledge about what to consider as evaluations or not. However, a voluntary omission from the evaluator allowed the respondents to freely discuss internal evaluations, i.e., the program's internal processes, and external evaluations, i.e., everything that does not revolve around the internal ways of working for the program. This led the interviewees to open the conversation on evaluations through their eyes and own experiences on evaluations and what they are. Following this, the interviewer could begin the conversations in either direction before regrouping to get the most information.

*"... next week we will have uh developing days, and we have planned to fill in the self-evaluation form about this program." (Interviewee 1)*

*"The steering group is the place where I can, like, evaluate, and the bosses evaluate me" (Interviewee 2)*

*"Students interviewed some customers and [...] they gave us the results" (Interviewee 3)*

*"Last year [...] we did do an internal evaluation." (Interviewee 6)*

*"One evaluation was done by university students..." (Interviewee 8)*

When asked about the importance of evaluation in a nonprofit context, all the respondents expressed high interest and mentioned the need for evaluation and getting answers on aspects of their work, such as the services, staff knowledge, or overall programs. Some questions mentioned as the ones in often need of answers through an evaluation were, for instance, whether the program is running as it should and reaching the set-upon goals, if the people within the program are offering adapted help and guidance to the customers, and members, or if the offered services correspond well to the need of the target audience. So, it is easy to conclude that the overall view on evaluation is positive.

*“Evaluation is always good. It shows what need you have [and] it tells us where we need to improve; it might also show where we are bling about ourselves.” (Interviewee 1)*

*“Evaluations are lovely, and I think they are really important for, like, this kind of service that we are doing.” (Interviewee 4)*

*“I think evaluations are valuable because there are results, what we do. Having the results is tangible; this is what has been done, what has been great, and what the problem areas and pain points are. So, you know what to address and what to celebrate.” (Interviewee 6)*

*“Well, I think evaluations are important because then you understand where you stand as an organization. It is a way to know that you need to re-strategize.” (Interviewee 8)*

Nevertheless, coming as a clear contradiction to that statement and a positive view on the importance of having an evaluation done, the experience of conducting an evaluation is not always positive. Some interviewees mentioned that they had previously conducted evaluations, internally or with the help of external parties such as students or non-professional evaluators, but mentions were made that the evaluations failed to meet the needs of the nonprofits. Indeed, the outcomes were seen as fragile and uncertain by the organization that accepted the evaluation, and thus, it did not reach the intended uses the internal stakeholders wanted to make of the evaluation.

*“Students interviewed some customers, it was only about the feedback of the customers, they made a presentation, and they only gave us the results once the satisfaction survey was completed.” (Interviewee 3)*

*“... we just evaluated the uh local impact [...] at the end of 2022.” (Interviewee 6)*

*“University students did one evaluation to observe how we do operations and uh increase participation, [...] honestly when I was reading through the report, it was just statistics, there was no outcome, it was not clear, it was all over the place. And with those surveys, it depends on the questions, who you ask, and how many answers you get.” (Interviewee 8)*

On the other hand, another part of the respondents answered that they had not seen evaluations conducted at all amidst the need and wished to get answers. For some, this is justified by the recent opening of the service (less than a year). However, for others,

observations on unspoken communications when discussing this topic demonstrated remorse and disappointment at the fact that evaluations are not being done although the need is present.

*“Well, no, I think no. We haven’t because everything is like we are just starting.” (Interviewee 2)*

*“... not yet, because we just opened, but we already have some numbers...” (Interviewee 5)*

*“Uh, I don’t recall seeing and evaluation.” (Interviewee 6)*

*“I think that organizations like ours [...] we are not evaluating enough. I think we should evaluate more; we would learn much more.” (Interviewee 7)*

The positive views on evaluations cannot be denied; it is clear that the respondents positively viewed the impact of evaluations in both internal and external aspects. However, negative experiences or the lack of communication in evaluations seen in nonprofits where the respondents currently work or volunteer cannot be omitted.

## **4.2 Program evaluation processes**

### **4.2.1 The choice of evaluator**

Discussing the importance of the evaluator with the interviewees of both public and volunteer organizations sparked interesting conversations with corroborated and contradictory opinions. Interestingly, respondents’ answers voiced similar opinions regarding skills, knowledge, and experience the evaluator must have to evaluate the use nonprofits want to make of it. However, they also conveyed two opposite sides regarding the type of evaluator who should be conducting the evaluation.

Regardless of the type of evaluator conducting, most respondents agreed that evaluating is not in everybody’s power. According to the respondents, an evaluator should possess

specific capabilities, previous experiences, and knowledge about the sector, either public organization processes or volunteer organizations.

*“I think it would be a really good option to do some kind of evaluation, but they have to have some background or knowledge.” (Interviewee 2)*

*“... the company who might do this evaluation, they have to have uh information about immigrants, the law, integration, and so on. Because it helps to understand the service and the whole process.” (Interviewee 3)*

*“I would like someone who has done something similar, worked in international houses, or developed these services elsewhere. Someone who knows. He needs to ask the important questions so that the knowledge would be good.” (Interviewee 4)*

*“... it is a little complicated in the public sector; you have to make a tender to make someone do the evaluation, kind of a selection of the evaluation. You cannot just choose, so it is easier to create it yourself.” (Interviewee 5)*

*“I would want someone who has experience [...] someone who can sport certain factors in qualitative data or someone good with excel or math can easily analyze these things. So, the evaluator needs practical skills” (Interviewee 6)*

*“Well, I think the person must have much knowledge about our organization, and uh, for instance, if you evaluate our courses, you should be active and know what we are doing there in the courses, the goals, and the target groups.” (Interviewee 7)*

*“I would like the evaluator to understand volunteering. I can imagine that if you have never volunteered before, you don’t understand the point, why it is important, and the impact it can have on the community. So, I would like someone with that understanding because that would make the evaluation much more fruitful.” (Interviewee 8)*

Although it is obvious through the answers that skills, knowledge, and experience are necessary qualities for evaluators, some interviewees preferred an external evaluator, while others preferred an internal one. Each respondent provided clear arguments as to why their choice was adequate. The respondents wishing an external evaluator justified their choice with the potential biases and blindness someone within the team may have.

*“I see the evaluator as someone who is not from [...] because for it to be like, uh, valid, proper, not biased, I think it would have to be an external person”. (Interviewee 8)*

*“Externally, definitely, because we can do the self-evaluation internally, but it is good to have someone outside because he or she sees our blind spots. External people are also neutral; inside, you might be biased or have an interest.” (Interviewee 1).*

*“... if you are from the organization, you don’t see, you are blind to these things” (Interviewee 3)*

The respondents wishing an internal evaluator legitimized their answer by the extensive knowledge about the nonprofit, the ability to know subtle ongoing, and a better understanding of the usability of a team member's evaluation against an external person.

*“Well, maybe we would do it by ourselves, create a team, and think about what we want to know and from whom we want to get the feedback. [...] I have noticed that outsiders often don’t know what we are doing, so it is harder to ask the good questions. So, we know better what we do, what we need, and what to ask.” (Interviewee 5)*

*“It’s always easier internally [...]. An external person would do much work over the quarter data, like if they have numbers or something, whereas an internal person would still know the subtleties of what is happening. They would be able to add that extra flavor to the assessment.” (Interviewee 6)*

*“Someone internal would be interested in doing it and also would see the use of it. You have to know what you want to evaluate, so that is, I think, better for someone who has worked in the organization to do the evaluation.” (Interviewee 7)*

However, one respondent mentioned the views of both parties. This person expressed the plausibility of biases and the advantages or disadvantages of choosing one option.

*“I think someone internal could do it, but not me because it is also an evaluation of my job, so I would be biased and not objective. Sometimes, it can be internal, but external is also a good option. Of course, it takes more money if it is external or someone’s time if it is internal. There are pros and cons to both aspects.” (Interviewee 2).*

Without a doubt, the majority of the respondents felt strongly about only one type of evaluator. However, all agreed that evaluators must be skilled and understand the context in which the nonprofits operate, alongside its agenda, strategy, and mission, to provide valid and usable answers for the nonprofits.

#### 4.2.2 Defining a program

As discussed in the literature, when evaluating a program, the definition enters the equation as it may change due to the evaluation process, thus finding its place in the final framework. Respondents were asked how they define their program and services for the tool to be usable. The answers on the topic provided rich strategic options from which interviewees usually chose when defining their program. Overall, the respondents agreed that the program definition remains in strategic positioning, with answers linked to environment, culture, purpose, or capabilities. However, it is worth mentioning that several interviewees also communicated the possibility of combining various strategic positioning factors to create the definition and emphasized through examples and words, given that the evolution of program definition is highly dependent on the experience within the nonprofit.

*“I use the mission statement, so I say that we are a service [...] that helps to integrate.” (Interviewee 1)*

*“Well, we have a slogan which I can use [...], and that was the moto, or slogan, which was already made before I started the job.” (Interviewee 2)*

*“I think I would, uh, I think it’s way heavier on the mission statement. [...] I think that is something we’ve been taking into consideration. But for me also personally, the goals matter that I have been planning in my portfolio, but I think most of what we have been doing, for as long as I have been part of [...], is quite uh. We’ve been doing it mostly to support the mission statement.” (Interviewee 6)*

*“If I would like to describe our organization and how we think, it would be the strategy plan. So, what we are doing. But first, you have to have a vision. The strategy would come from our vision.” (Interviewee 7)*

*“When we first started, it was my statement because [nonprofit name] was just created, and it was the reason why [nonprofit name] came to be in the first place. That statement talks about my struggle as an international person. [...] Then it started to be built, and now we have a vision [...]and a mission [...]. So, to answer the question, it is a combination of everything, but recently, with our new strategy, it comes from hearing the people who come to our events, listening to what they say.” (Interviewee 8)*

The program definition differs according to the decision-maker being interviewed. Thankfully, reflections are given on why that is, and several factors can be identified from

the answers, such as the resources and experiences available to the person describing the program and how long the person has worked within the nonprofit. However, the consensus in the answers gathered from the respondents remains very closely linked to strategic positioning factors.

#### **4.2.3 Conditions and actions**

The conversations with the respondents took the longest when discussing the program evaluation processes, as much information was released during this part. Two conditional categories appeared to be recognized through the interviewees' answers: internal and external. Both conditions would impact the nonprofit through the program's different actions. Hereafter, the paragraph explains the conditions for voluntarily omitting the separation made in the interview process, as respondents often did. Such writing will allow the reader to quickly understand which conditions impact which actions. For clarity in this dissertation, the conversation will first revolve around internal conditions followed by external conditions. However, the reader is advised that during the interviews, the conditions and actions were discussed in random order simply depending on the interviewees' chain of thoughts and experiences sharing.

The internal conditions discussed during the interviews concluded in a list of five conditional factors that impact the programs' actions. The factors identified by the respondents are statistics and data, cultural and field knowledge, human resources, time, and capital resources. However, not all internal factors were given the same importance by the respondents. Some were mentioned to be an absolute necessity, others a mere information tool; similarly, some internal factors were popularly mentioned by all respondents, or the majority, while others were only expressed by a few.

Thinking about the statistics and data was only a matter for a quarter of the respondents, who expressed that the numerical information regarding their services is important to justify not only the need for the program to keep running and existing but also to know

in which direction to put the resources in the future, as this depends on the customers' needs and expectations.

*"Also, what is coming to my mind what we need [...] is the statistics; how many people are coming, what are their questions, how we helped them [...]. It shows that we need permanent work and more resources [...], it's good to have such information because we need the knowledge to develop the program and the service." (Interviewee 1)*

*"... which nationality, which languages we used, how long was the meeting, was it in the office or by phone, or by written contact, where the customer heard about us, [...], gender, age, country of origin [...]? And decision-makers usually don't know too much about immigrant work, so for them, the numbers and facts are more to show that this service is needed." (Interviewee 3)*

Interviewees also mentioned the importance of the knowledge needed when working in the nonprofit sector dedicated to working with internationals in Finland. More details on cultural and field knowledge were mentioned. Knowing the culture is essential, but so is knowing how to work with immigration, staying updated with training, and having the answers to questions or knowing where to find the answers. Here again, only a few of the respondents expressed that the need to know about the field and the culture is a condition that impacts their actions.

*"I was in training in Lahti, and it was good to get to know the people. So, it has been useful to know how they are doing this and this, what system they are using, and so on." (Interviewee 1)*

*"Cultural knowledge is important. Also, we participate in some webinars [...], for instance, about immigration in Finland, and there are many events every week, a meeting where you can meet and listen to the subject that you are interested in. [...] This is where I can also learn additional information that we can then advise well and correctly to our customers." (Interviewee 3)*

*"The knowledge is needed, at least the knowledge where to like, look for information, I think. We don't have to know exactly everything, but you have to know where to find the information. That's the thing, and that's what we do; we find the information or the place that can answer the question..." (Interviewee 4)*

Indeed, both conditions above are important for the respondents who discussed them, but the reader is informed that the answers originate from one type of nonprofit. They are all respondents from public organizations. As the study considered public and

volunteer organizations for the research, the aspects will be placed in the final framework. However, it is crucial to note that considering statistics and knowledge as important internal conditions relies on public organizations only.

Nevertheless, several internal conditions were as important for volunteer organizations as for public organizations. Those conditions are human, time, and capital resources. Indeed, each of these is linked to similar actions conducted by the program, and lacking one or the other may drastically reduce the program's actions. Firstly, let's dive into the human resource conditions, which were mentioned by all but one of the respondents as a key condition for their program and services to run smoothly. Interviewees explained that having people understand the customers' needs and conduct the program's events and services is mandatory; nothing would be done without people.

*"We need people because we can plan a lot of nice things and many promises that we will do this and this, but if we don't have people working towards that goal, we have only empty promises. So, I think that is the main thing: we need people capable of customer service and others who can be flexible in their jobs. [...] For example, we can do much marketing, but it does not make sense if we don't have the people to help the customers." (Interviewee 2)*

*"... before it was only two people working, and now we got one more working here. [...], but [name] was just hired now and for only a certain time and then after we are back two people [...]." (Interviewee 3)*

*"... so that people they don't have to wait weeks. Like that there are enough people to be present and be like ready to help so people can walk in some days of the weeks and get help." (Interviewee 4)*

*"Uh well, all who are involved there should be kind of uh motivated participants, because the program it does not run by itself." (Interviewee 5)*

*"So, I think uh, we need people [...] we need proactive people and uh those people also have to be quite focused on their portfolios and what needs to be done for the portfolio [...] so I think, internally, that resources need is the people. And also, volunteers mean events." (Interviewee 6)*

*"I would like to step aside and have others take over, but it is difficult for others to know how Finnish organizations work. And, uh, well, we need and have board members responsible for certain things." (Interviewee 7)*

*“We need resources. And what I mean by that is both human and capital. So, need human resources, you know, like [...] volunteers [...] our action is our events, so to coordinate it from scratch to plan it, it links because we need people to host the events.” (Interviewee 8)*

When discussing the human resource factors, some respondents also made the connection with the time. Time is also a recognized internal factor, which was mentioned by fifty percent of the respondents. According to the answers, the time nonprofit volunteers and workers grant to their actions within the program is a core internal factor. Indeed, time is a rare resource that people need to give to nonprofits for it to run.

*“Umm, like time resources, that we are available, so people don’t wait...” (Interviewee 4)*

*“... because we are all volunteers, we cannot be constantly getting into others’ portfolios to do their job, so we have to work synergistically with each other...” (Interviewee 6)*

*“Finnish lifestyle is hectic, and people have a lot to do, and organization is not that popular anymore. Now [...] I have time, but people who work they don’t have time anymore.” (Interviewee 7)*

*“...because we are using our time, which is valuable and sometimes hard to get. When you are volunteering, it can be hard with our normal daily schedule, especially since we have families and we have work.” (Interviewee 8)*

Finally, capital resources were an internal factor mentioned by half of the respondents as a strong condition to consider for the program to conduct its expected actions. Without it, actions are difficult or close to impossible. Answers mentioned that having a budget helps the nonprofit run daily operations, such as maintaining a website, and bigger actions, like conducting events or providing good customer service.

*“The form is very good because then we can justify the permanent need for the service, and also if we saved money or spent some money, if we used translators or not. Then we can say we need resources.” (Interviewee 1)*

*“Well, we try to find different ways [...] now, for instance, we are struggling with the resources, we have been trying to be innovative to find new ways that even though we don’t have the biggest resources, we have a way to make the customer happy. It is challenging sometimes.” (Interviewee 2)*

*“Even for the smallest thing, we constantly need money; the money is needed, for instance, for flyers. [...] Having funds helps in many aspects. Internally, for example, I mean more on a psychological level [nonprofit name] as a motivation tool because everyone knows they would get paid to do something.” (Interviewee 6)*

*“... and then, of course, we need capital, we need money to run the operations, like the website, it’s not super expensive, but it needs money, and we need money to pay for it. So, umm, I would say those two things are what is needed to run. [...] We have certain events we may have to pay for. Usually, our events, we tried not to have people pay, or not have to pay upfront [...] bit sometimes we need to pay for certain activities to have the event.” (Interviewee 8)*

The conversation on internal factors resulted in a great list of interconnected conditions and actions thanks to which the programs can run and exist or without which difficult situations would be implied. Certain factors are solely important for one type of non-profit, while others greatly impact both volunteer and public organizations. Overall, the respondents openly disclosed that certain conditions are necessary to survive. Alternatively, the external conditions necessary for the program to run its actions gathered a smaller list with three main factors identified through the respondents' answers. The factors are human resources, networking and communications, and partnerships; once again, each factor guides the completion of various actions.

Although the human resource factor was discussed in the internal section as necessary, external human resources were debated as customers and members. However, an important point is that only half of the people mentioned customers and members as conditions required to run the program. Those interviewees said that without customers or members, there would not be any need for their programs and services, leading customers to constitute an external human resources condition needed for the program to carry out its actions.

*“... but the customers are always the most important part, so I try to always think about them first.” (Interviewee 2)*

*“... and also, we need to have uh... customers.” (Interviewee 3)*

*“Oh, and also, it is the people externally because without the people, [nonprofit name] wouldn’t be there. So, people that are possible members. Without people on both ends, there is no [nonprofit name].” (Interviewee 6)*

*“We have to think of our... Whom we are expecting, whom we are serving and who are the clients. [...] Uh and we have to think about their needs.” (Interviewees 7)*

However, no other interviewees mentioned customers and members as conditions. Surprisingly, all the respondents mentioned customers and members in answers to other questions. Customers are important but were forgotten as a condition necessary for running the program. Hypotheses could have been made on the hidden meaning of some respondents’ answers but were not for reliability purposes. This brings contradictory information to the study as the customers are central actors in all the respondents’ answers except for the question on necessary conditions to run actions. So, on an external level, human resources are also mandatory conditions for the program to run, and they will be represented in the final framework. However, the reader is advised that in the final framework, links are made without noting the difference between the actions resulting from internal or external human resources. This is justified as internal and external conditions link to actions defined with similar names, although the outcome is different; thus, the analysis of the results details the differences, but the framework solely presents the possible actions of the human resources condition.

Two conditions closely linked to each other in the respondents’ answers were networking, communications, and partnerships. Indeed, most respondents agreed that one or both are conditions for the program to conduct a given set of actions. First, networking and communication are needed to reach the intended audience and execute events and services. Thanks to networking and communications, nonprofit organizations can run events and present their services, but they can also use their network to collaborate on some activities or use their network to assess the customers’ needs in terms of activities or answers needed.

*“We need local network, national network, education [...] in every aspect—at least all the basics about needs and immigration in Finland. We also meet with other*

*immigration services from [city] and the region, where we all network and discuss the tasks and services. We all share.” (Interviewee 1)*

*“We need to participate in webinars and training [...] not our own but participating in those of the city, or where we present our service. Also, we can attend events with our partners, governmental and national meetings and seminars, and networking events with other immigration services; it is continuous learning. [...] Also, we do benchmarking and learn how they do and what their services are from the other people.” (Interviewee 3)*

*“... uh, and maybe pop-ups, for example, with Kela or Migri. We need many actors who work together and find our concept important. We have managed to do it well because different units and organizations want to work with us.” (Interviewee 5)*

*“We need communications, and we need to find the right communication channels where we can find the members. That’s perhaps the first thing that is coming in my mind about necessary conditions.” (Interviewee 7)*

*“For some events, not for all, then it is necessary conditions. For example, what I am thinking about is [event] because it is tonight, then we need to have a place to host it, so to have the place then, we need permission from the venue, so in that sense, it is needed...” (Interviewee 8)*

The discussion around necessary conditions ended here for some respondents, but others mentioned the need for partnerships or briefly explained the role partners may play in the actions led by the programs.

*“The activities that are done correlate with the partners that they have. They find value in being part of our activities. They can drive new customers and also certain information. It’s both ways a win-win. We work closely with [name of a partner] and organize joint events to promote our services.” (Interviewee 5)*

*“... We need partnerships, not just regular but also more strategic ones. And uh, I think because even if we think about reaching a certain audience, we still would need cooperation, so I would uh certainly have [...] partnerships as [an] external conditions needed.” (Interviewee 6)*

*“... but for example, our [...] events, you know it’s great we have a partnership then, it’s great they give our [nonprofit name] members a discount, but we don’t need their permission for us to be there, we can pay the full price and just go.” (Interviewee 8)*

Understandably, external conditions are needed for the program to involve either user of the programs or third-party providers, which is valid for both public and volunteer organizations interviewed for this study. Additionally, half or more of the respondents corroborated the three conditions identified, providing good reliability in the answers.

Discussing the program evaluation processes sometimes left the interviewees uncertain of their knowledge and whether their answers would help the study; spoken and unspoken communication somewhat confirmed this hypothesis, as gestures and filling words were used when answering questions on necessary conditions. However, thanks to the answers and experience sharing, this section on conditions and actions can conclude by stating that one way or another, both internal and external conditions, via sub-categories named by the respondents, are actively touching the program's actions, positively or negatively.

#### **4.2.4 Context**

For the study, the respondents were asked to divulge information on internal and external contextual factors that can impact the program. Unlike the conditions discussed previously, which are necessary for the program to run its activities, contextual factors are not mandatory must-haves to conduct actions; they are simply factors that can impact the nonprofits via different forms, positively or negatively. The core difference here resides in the prerequisites that conditions are and contextual factors or not. The results analysis will go through internal and external factors, starting with the internal contextual factors.

When discussing the internal factors, respondents mentioned similar factors as for the necessary conditions. Amongst the answers to internal contextual factors are human resources and capital resources. Several respondents discussed human resources because it is a condition required to run the program and an internal factor that can impact the nonprofit in some ways.

*“... so because of that, maybe we have fewer people and then have to do other tasks...” (Interviewee 1)*

*“I think with the good amount of human resources as we can, it makes everything easier. Because it is from people to people. As long as we are working with people, we need people to work with them.” (Interviewee 2)*

*“Absolutely our resources, that’s a definitely yes because we would need more.” (Interviewee 5)*

*“Uh, I think just the people like volunteers who agree to work. People matter for any association to run because we are often more immigrants than local [...]. The people matter the most, so we need to focus on the people.” (Interviewee 6)*

*“You may never know if there are no volunteers or board members. Hopefully, when no human resources are available, it will never happen, but then [nonprofit name] would not exist because we would not have anybody running or hosting. Um, so, I think that is the biggest thing.” (Interviewee 8)*

Additionally, the budget is a recognized factor that can impact the program directly. However, one difference is noted in the answers. For some respondents, the budget can have a direct and drastic impact on the program, while for others, the budget is undoubtedly a contextual factor to consider, but not one that would dangerously impact the program.

*“money for sure because we had a budget cut. The future is open now, and we don’t know what will happen then. [...] So, the development will be slow then, we have to make it slow because I have other tasks if we have fewer people.” (Interviewee 1)*

*“if money ends. Budget yes because now you make budget forecast for next year, but they already cut a third of the budget. So that means for next year we have less budget.” (Interviewee 3)*

*“And yes, the resources, especially the money. Now, it’s really like what we see every day. I don’t know if I will work there after February, so What can we do next year? How much can be done?” (Interviewee 4)*

*“As said, we need more people to work at the info desk, but it depends greatly on the city budget. [...] We get the budget from [...] and also from [...], and they are selected yearly and also salaries are decided. They are different budgets to run. And also, how important actors inside [city] find [nonprofit name], if they don’t find it important, then they don’t give the budget.” (Interviewee 5)*

*“... of course, it would affect us, but it would not be a huge drain if we did not get grants for a certain time. Because you know we have three founders, so we help in the investment part, so that is not an issue, especially since, as you know, our annual costs are [...] very, very small, so it is not a huge deal, and I know that in the future we will be able to find a way to get fundings.” (Interviewee 8)*

One respondent also mentioned that time is an internal factor; however, due to the lack of other mentions during interviews, this factor has not been included in the final framework. Yet, the reader must be aware of such mentions.

*“Uh, very practical things like how much time we have. We are so few in our board, and we are the ones who are acting, so uh yeah, these kind of things like time is also dictating what you can do.” (Interviewee 7)*

As nonprofits, it is obvious that capital and human resources are mentioned several times in terms of different aspects of the program. Indeed, such resources appear to the respondents as rare yet important for the program runners. As far as internal contextual factors go, none of the respondents could think of aspects other than these two or three that could have an impact, positive or negative, on their program.

On the other hand, the interviewees had a lot to share when discussing the external factors. Due to recent events that touched Europe and Finland, the respondents had hands-on and recent experiences with what outside factors can directly affect the program. Of course, one of the first factors that came to mind for the interviewees was related to immigration numbers. Factually, the number of immigrants entering the country can highly impact the programs dealing with internationals living in Finland. Some respondents mentioned planned immigration, such as student exchanges.

*“Uh, what also impacts is the number of international students, because now the services are full, and they are panicking like what to do.” (Interviewee 1)*

*“The students also said that they don't know what to do if they come and everything is full. If they come with their family, by law, they have to offer that if their child is from six to eighteen, the law says they must be in education. They must be, but if there is no place, then this impacts.” (Interviewee 3)*

*“.. how many like immigrants are moving to [city]. That’s connected to education, looking how many students they will get here, and how much information they give to their students and their families.” (Interviewee 4)*

*“And what else, maybe how many people move here, and how many students the university has? Also, some students are struggling to find jobs, so they contacted us. So, we have noticed that also.” (Interviewee 5)*

Another section of the respondents mentioned immigration numbers as unexpected immigration numbers, for instance, refugees entering Finland due to wars.

*“If there comes some new crisis, for example. For example, we have a lot of Ukrainian customers and guide them. So, if another crisis comes in another country and we get people from that country here, that impacts.” (Interviewee 2)*

*“For example, this uh, Ukrainian war, [...], now because it came more people.” (Interviewee 3)*

*“Having uh more immigrants coming to the country works great because that is our target audience, and [...], yes, we are trying to bridge the gap more than the locals; it is the internationals who need, you know, to feel that they belong somewhere, so they are a factor. So, increasing or decreasing those numbers would impact us.” (Interviewee 6)*

Furthermore, the respondents mentioned the governing bodies, both local and national. Indeed, Finland, the country where this study was conducted, has recently elected a new national government that is rethinking the bills in place granting citizenships, permanent residence rights, and other crucial aspects that immigrants must deal with regularly. As such, the programs and services dealing with international people daily are impacted by the potential changes and the stress these may cause to their customers and members.

*“Also, when this voting for the election was last spring, more of the Finnish right party came to the city, so that can change our program. [...] Also well, from June last year, there started to be contacts and people come who were not yet worrying about applying for citizenship, or with the permanence residence permit, and they want now to apply as early as possible because they don’t know what will happen next year.” (Interviewee 3)*

*“Definitely laws, also like how the integration and everything around it works.” (Interviewee 4)*

*“Politics, for instance, we know this when the new government was selected; there were many uncertainties from internationals living here and what would happen. There were many questions for us, but we could not answer them because we did not know. Maybe they did not understand the Finnish political system, and the matter had to be voted on. So, politics definitely, and in many cases, they affect negatively.” (Interviewee 5)*

*“Also, uh, certain government policies are at the city level. For example, if there are more targeted towards immigrants, certain grants would be available for us to pitch for. But, uh, if things go against immigrants, there will be fewer grants for us, which means it impacts us like that. Also, this whole government thing will impact local sentiment, so the local feels like the government doesn’t support the immigrants; they might feel biased towards that and might be like, “... we should also really think about this,” and that you know if the government as a whole is not supporting then maybe we shouldn’t. So this also would have some impact, I don’t know how negatively, but it would certainly have an impact.” (Interviewee 6)*

*“First thing I think about is politics, like if something were to happen like this new government and there is like a huge backlash towards immigrants like that’s going to affect the way [association name] operates because we focus on immigration, we focus on these internationals immigrating into Finland. So, I think that’s a big factor, as is immigration policy.” (Interviewee 8)*

Nonprofit organizations consider political and immigration agendas to be highly sensitive factors. A slight change in one or the other could completely derail their program and services, or it could also mean a positive change as the need would increase, providing that the program would be adapted.

Aside from the political and immigration factors, others have been mentioned, amongst which partnerships are taking third place. Indeed, a clear connection can be established between the people reaching the nonprofits interviewed and the level of information granted by third-party providers to those same customers. The answers can also show extra links, such as collaborative work and investigating new ideas and needs.

*“Well, our point is to combine the projects and organizations helping immigrants and doing things for immigrants. We want that [nonprofit name] to be the one who has them together and makes the cooperation better. For example, we organize project meetings where people like to represent new projects; it’s networking. We want to be the ones who know, not only for the customers but also for those who also help and guide the immigrants.” (Interviewee 2)*

*“Of course, how well is the service in the city or nearby? I think everything is about how willing people are to serve in English. So, if there is good service that they can serve their customers, they don’t need to come to us.” (Interviewee 3)*

*“Yes, if the university cannot help their students [...], they come to us even more. Uh, I think the most important service nationwide is the kind of service where you can call and ask [...], and if they would shut this down, then there would be consequences because we don’t have resources.” (Interviewee 5)*

*“Uh, we are cooperating a lot with the city regarding what they are developing. Tomorrow, for instance, we have a meeting with all the actors working with immigrants. So, these external factors give us some idea of what needs to be done, for instance.” (Interviewee 7)*

The conversation ended there for most respondents, but a few also mentioned customers as an external factor. Indeed, when customers do not need the service anymore, they are not reaching out, meaning that the nonprofit will be impacted; on the other hand, more people mean more customers and more work.

*“Well, I think more and more people now know that [nonprofit name] exists, and if [...] externally people would know more and more about [nonprofit name], then we would have more and more customers. But then we cannot control the number of customers, which is also a positive problem.” (Interviewee 2)*

*“... now we have more work. [...] Of course, this to us means more customers.” (Interviewee 3)*

*“This is a positive thing, but if it is not needed anymore. That is the goal. To be in a country where [nonprofit name] and associations that help people integrate is not needed. That would be a factor. It’s all I can think of right now.” (Interviewee 8)*

Although these were the most common answers, it is important to note that four other factors were mentioned by a very small portion of the respondents. One factor was disclosed by twenty-five percent, i.e., two persons, while others were individually mentioned during different interviews. The first factor was the location of the nonprofit. If the organization had a dedicated place to welcome its customers, this could be considered something with the power to impact the program.

*“... also, the space, or location. Because now we are in the center and the space is easier to find, so it had an impact on the service.” (Interviewee 4)*

The second factor was the reviews or opinions that members and customers have and share with others regarding the nonprofits. Sharing positive views would positively impact the program, while negative reviews may lead to fewer customers using the service.

*“But a thing is also that if you get a bad or any services, then I think it is easier to tell. The bad comments spread more easily, so if you once did something negative to the eye of the customer, then what are the chances that the person will come back? So, then, it is hard to get the trust back.” (Interviewee 2)*

*“Uuh. Members opinion, members like what they are expressing.” (Interviewee 7)*

The third factor was the program's status, whether permanent or temporary, as the attention received and the power of the nonprofits is not the same.

*“The status of the program. Because we are now permanent, [partners] invite us to their groups. [...] Being permanent helps in many ways. For many network organizations, for example, it is easier because they like to have a permanent service where they can guide the people directly.” (Interviewee 3).*

The fourth factor is that competitive services or similar volunteer organizations could impact their program.

*“Other competitors, I would think. Even though I feel like [nonprofit name] has a very good position, um, because many nonprofits focus on a niche, and we have our niche.” (Interviewee 8)*

As each of the four factors above was only mentioned once or twice, there is not enough data to confirm that these three factors are indeed contextual matters related to the nonprofit context, and thus, they could not be integrated into the framework. Further case studies would be needed to confirm each statement's validity.

Overall, the conversations on contextual factors were two-sided. Discussing the internal contextual factors surrounding nonprofits created very similar talks between all the respondents, with clearly three resources as a core focus. However, the external factors provided a larger base of answers, with individuals reflecting on numerous factors, some backing up and others not. Yet, the reader is advised that recent events in the country happened during the interviews. Thus, they may have influenced the interviewees' answers regarding the external factors.

#### 4.2.5 Users and uses

Discussing the users and the uses of evaluation was a stressful topic for most respondents, as worries about the lack of knowledge or uncertainty in the answers were expressed even before the start of the interviews; unspoken communication confirmed part of those emotions during the interviews. Notable is that this section required updates in the questions' design during the questionnaire testing phase to facilitate the conversation and minimize the tension and unsure feelings of the respondents. For clarity, it is noteworthy to remember that the respondents were questioned around several theoretically relevant themes within the program evaluation scheme to provide interesting data regarding users and uses of evaluation. To inform on users of the evaluation, the questions revolved around *who*, so, who would participate and who would access the results. The questions around *what* aimed at gathering data on the uses, what use of evaluation, and what section of the evaluation. One question gathered both sides to link the *who* and the *what* together. The reader is invited to review questions eight to twelve in Appendix 1.

This section will disclose each of the four user groups identified and their corresponding uses. In doing so, the reader will more easily understand how the respondents linked one user group to its supposed or imagined evaluation uses. The respondents have identified both internal and external stakeholders. The analysis will present the classifications in the order in which the nonprofits mentioned them. Surprisingly, most nonprofits have mentioned the users similarly, giving the reader a chance to discover the users as presented to the researcher in the interview process.

The customers were the first group identified by the respondents and were mentioned by everyone interviewed. In the first part of the conversation surrounding this user group, some interviewees disclosed that the customers would be an interesting addition to the evaluation process, meaning that they see the customers as having an active role in the evaluation rather than passive. The customers are here seen as a means to bring added value to the evaluation and tailor the use of the results toward the needs of the nonprofit.

*"... also, the customers because they are the ones using the service, so we need to know what they are thinking, what is the services they need, don't need, etc., we have to make sure [...] that it is what they need and know what they are saying." (Interviewee 1)*

*"So, I would like to have the customers there. Because I think, in my head, like the customers they evaluate every day, if they use the services more than once, it gives me a signal that they got a good service on their first visit. So then, if we get customers like that, they are coming... Then we can be, in a way, satisfied that we see that people are coming and also in our events. [...] Customers are the best to evaluate if we have been doing what we promised." (Interviewee 2)*

*"... to us, clients are the number one priority." (Interviewee 3)*

*"Um, the customers, for sure. [...] It's like the customers know if they got any information from us because it is all about them. That's why it's interesting to know what they think about our services and what they feel." (Interviewee 4)*

*"Also have to ask the clients, those who come to [nonprofit name], what do they need in the future?" (Interviewee 5)*

*"Whatever evaluation we have done internally has been purely towards the people that [nonprofit] has impacted." (Interviewee 6)*

*"Uh, of course, I would like every member of the association and people that have met us [...]. I want to know what they think about our actions and what they want us to develop or change. Because this is very important, how would we know [...] if we don't evaluate. Of course, we should ask before we create something, but afterward, it is really important to have their opinion." (Interviewee 7)*

*"Well, I would think, you know, certain members of [nonprofit name] because they need to explain how we work, and for whoever is doing the evaluation, they need to understand. As you know, every association and every nonprofit is different, strategies are different, and our missions are different, so we have different values and ways of working. So maybe one or two people from [nonprofit name] will support the evaluation process." (Interviewee 8)*

Additionally, the customers are linked by several respondents with the outcome use of the evaluation, meaning that the customers would use the results and conclusions. Unlike the outcome use by the internal stakeholders discussed later in this analysis, the outcome use for the customers is more passive. The respondents mentioned that the customers are incentivized to analyze the results and conclusions of the evaluation in

which they have taken part. It is remembered here that the respondents have placed the customers as actors within the evaluation process.

*“But maybe the customers also, this is a public service, so maybe also the public could know about the results.” (Interviewee 1)*

*“Also, maybe we could share with the customers who are interested.” (Interviewee 4)*

*“For example, I think uh someone who spends money, like as a member, for example, if I were a member, I would want to know what is happening because I support with whatever amount in the organization so uh, I would want to understand how they are running it.” (Interviewee 6)*

*“Like every member. I want to publish the results publicly on our local network. Of course, on Facebook on our website when it is ready. Yeah, the results are not a secret for our organization.” (Interviewee 7)*

One interviewee mentioned that clients and members might also symbolically use the evaluation. By observing that an evaluation is conducted, they might feel more inclined to safety and comfort, knowing that the program is ensuring their smooth operations and strategy. However, due to the lack of mention from other respondents and the slight risk of misinterpretation on behalf of the interviewer, the connection is omitted in the final framework. Yet, the reader should be aware of that mention.

*“...maybe the customers would also be more like, uh, aware that we are planning and taking it more seriously. So yeah. Customers notice that we are thinking about the plans and adapting them to the customers, and then we get them to feel more comfortable and more like, uh, safe and comfortable.” (Interviewee 7)*

Customers are important for nonprofits; they are recognized as the target audience and customers of the program. Furthermore, the interviewees immediately mentioned them as a user group, and links to two different uses were made, highlighting the presence of the customers in the minds of the nonprofit decision-makers.

Secondly, all the respondents quickly pinpointed the following user group and four uses were linked to this group. Several actors were identified as internal stakeholders, the final framework will use this classification as a visual representation, yet the reader is

informed here again that this title includes the decision-makers, employees, and board members interviewed during the research process. About seventy-five percent of respondents mention the internal stakeholders as direct actors in the evaluation process.

*“Decision-makers, of course, are important because they give the money, so we have to hear them.” (Interviewee 1)*

*“But then, of course, the bosses and the directors [...], they are the ones who pay the salary...” (Interviewee 2)*

*“Definitely, the decision-makers, us.” (Interviewee 3)*

*“... and of course, also ourselves.” (Interviewee 5)*

*“So, we had internal people who did it, so I was doing the calculation, but we evaluated all sort of things together, I would say like mostly I was doing the calculation, and [hidden name] was helping with the data.” (Interviewee 6)*

*“I would say one board member and one volunteer at least.” (Interviewee 8)*

Additionally, all the respondents explained that they would use the evaluation as a strategic or accountability tool. This reflection on the results and conclusions of the evaluation is directly linked to the outcome of the evaluation. Indeed, according to the answers, the results and conclusions would be the section that informs the nonprofits the most on the goal set for the evaluation. For instance, some respondents mentioned it would inform them of the following steps, the improvement possibilities, or the needs of their audience. Such outcomes become the nonprofits' incentives to conduct the evaluation.

*“If we want to develop our system here, it is very important that at least we have the results.” (Interviewee 1)*

*“As an accountability tool. So, if it would be public, it would be like more motivating to try to fix the things that are shown in results.” (Interviewee 2)*

*“...it was useful because there was feedback from the customers and what the needs of the immigrants in [hidden city].” (Interviewee 3)*

*“Umm. Well, I think, of course, maybe we need the results and conclusions...” (Interviewee 4)*

*“To improve the activities, if there is something that is not working that well. So, definitely it would be important to know what to improve.” (Interviewee 5)*

*"I think it would definitely give us a better idea of how to approach certain things. Having this evaluation in place will make things quite process-driven, and I know how people hate some processes, but if we just have smaller things, it does not have to be seen as a process but just as something that helps us, so we are accountable. We are all spending precious time, [...], and discussing issues, but just discussing the issues is not the idea. We should also have solutions, and to find the solution, you need to have some process, so I would consider the evaluation as a process to dissolve our challenges and improve things in the future." (Interviewee 6)*

*"It would be the goal to know! We take for granted that we know what people want, but we want their opinions about their needs. And also, of course, how they experience what they have been offered so far. We want to adapt the activities more to reality." (Interviewee 7)*

*"I would use it for two things. Yes, the results are very important because it is the sole reason why we do these evaluations ..." (Interviewee 8)*

These two linked uses were the most popular among the respondents when discussing internal stakeholders; however, some of the participants in the research also mentioned two other uses that are interesting to discuss here. Indeed, fifty percent of the interviewees mentioned these uses as linked to the internal stakeholders' group, making the connection in the final framework relevant. The symbolic use of evaluation, discussed in the theoretical review, appeared amongst the respondents' answers as they expressed that sometimes conducting an evaluation is a means to an end. For grant purposes or justification for the nonprofit to exist, a few confirmed the symbolic use of the evaluation.

*"Evaluation also shows that the service is needed; it opens things to everybody, also to us." (Interviewee 1)*

*"... so that they understand what the need is because it was not permanent service yet, but it was a project. So, the results were used to prove the needs of our program." (Interviewee 3)*

*"We would use it to send it to [hidden city] who are like saying is this service good, or not." (Interviewee 4)*

*"But yes, when it comes like money or something, it is great to have it out there for the external stakeholders to see." (Interviewee 6)*

*"I would use it for two things. [...] and would also use it as something concrete to share with our granters, like you know, when we apply for grants. I would use it*

*since we are a nonprofit, and much of our funding comes from reporting these things. Show we did it and get grants.” (Interviewee 8)*

Last but not least, the methodology is an evaluation factor that several respondents also felt was important for internal stakeholders to use. As a line of improvement for future evaluation or as a guide to target a suitable audience, using the evaluation methodology is a critical criterion for several interviewees.

*“Maybe, but I’m thinking now more about, like, some people in steering groups would be more interested in the methodology, but in a way, everything has to be somehow in it. You cannot do like, of course, results, you can say, but how did you get the results? but you need the other things there as well.” (Interviewee 2)*

*“Well, I think also we need the process and the methodology maybe, like about the evaluation.” (Interviewee 4)*

*“Of course, the structure is important, [...] the methodologies too, we should think about that and how we collect data [...]. All are important, not only the results but also how we get there. After the first one, we would know how to improve the evaluation process for the next time and how to reach the people the best. We use the first evaluation to tailor the second one better.” (Interviewee 5)*

*“... we would have to think about the content of what we are doing and try to pick up uh clear structure of that content, and after that try to write down or to ask, talk with people ask the questions corresponding to the content.” (Interviewee 7)*

The internal stakeholders are the only evaluation users’ group identified with all the possible uses of the evaluation provided by the respondents. Overall, the interviewees considered that they have many different uses for the evaluation due to the multitude of themes an evaluation can have, according to their views on evaluation. Not only that but the use of the evaluation from the internal stakeholders’ perspective is also linked with the analysis of the strategic learning phase, which is coming up later in this research.

The third user group identified through the answers to the interview questions is the partners with whom the nonprofits often interact. Whether governmental or private companies or services partnering with the nonprofits, they have been included under the same umbrella. Not segregating the partners is justified, as the nonprofits do not do it themselves. Thus, grouping all partnerships under the same umbrella helps create a

more useful final framework. The link between conducting the evaluation and having the partners as users was made clearly by 5 respondents.

*“the governmental partners because you can come here and ask the stupid questions, “I don’t understand that, what is this?” (Interviewee 1)*

*“Partner services absolutely, like TE-office or Kela” (Interviewee 3)*

*“This is a hard question, but maybe partners ...” (Interviewee 4)*

*“Well, our partners for sure, so the university, multicultural center, chamber of commerce.” (Interviewee 5)*

*“The others would be, uh, our potential strategic partnerships. By strategic, I mean, for instance, companies that think about investing in us in some way. [...] I think the stakeholders would be mostly interested.” (Interviewee 6)*

And, for those respondents who identified the partners as a user group, the use these persons would make of the evaluation is very clear. Partners would only focus on either the outcomes or the evaluation methodologies. Thus, clear links were established between these two uses, while other evaluation uses were not linked to this group.

*“... you probably don’t ask in TE-office what this means, so they would know maybe what here people ask because it is easier here because [nonprofit name] is at the same level as the people.” (Interviewee 1)*

*“... they would be interested to know the activities that we do. For example, I was surprised because other services they send immigrants to us [...], the service helps the customers to understand. But because the law regulates other services and Moni-info is not regulated by law, we do what is needed to help; we have laws and lines that are not as strict as TE-service or Kela. It is easy also for customers to come to us and ask questions, rather than go to authority.” (Interviewee 3)*

*“... partners will know the information have we advise it right or wrong.” (Interviewee 4)*

*“I think the partners would mostly focus on the results. However, if they find good questions that they could use for themselves, they could also check the methodologies. But the results are the most important for them.” (Interviewee 5)*

*“So, I think for partnerships, yes, it is uh important to see the results, but it is also important to know the process, the methodologies, so I think every part of it would need to be looked at through a different lens. Not just the same thing every time.” (Interviewee 6)*

One of the respondents mentioned that partners could have a strategic role as actors within the evaluation, but without further data to confirm this fact, the option is hereafter presented to the reader, but the connection was not established in the final framework.

*“I think it will be useful, like, some officials, like working like TE-office or DVV or like some places that we guide people to. So, it would be useful to have them in the process.” (Interviewee 4)*

Not all the respondents have highlighted their partners as potential evaluation users. However, the data shows that this user group is identified in both types of nonprofits participating in the study; few respondents from both public and private organizations have identified partners as a relevant user group. The classification is named *partners* in the final framework, and the link to both identified uses is represented.

Finally, the fourth user group identified by three respondents can be defined as similar nonprofits. These possible users of an evaluation conducted by the nonprofits would be people who work in similar services within the same environment. For these respondents, the outcome would only be a strategic reasoning of identifying the results and conclusions of one study to adapt them to their programs. Although one respondent provided a more detailed idea regarding this user group, to think about a reporting system where all evaluation data would be gathered in one place for other similar nonprofit organizations to use the results of all evaluations shared openly.

*“Also, maybe other organizations, who are working in the same field, if they are like wanting to hear about [nonprofit name], or see what it is, then I think maybe this is a good thing.” (Interviewee 2)*

*“I think because all around Finland there are the International Houses in cities, or like similar services, so and I think every city is like doing a little bit differently the things, so it would be interesting to know how it is done here and they could use what we have done for them too.” (Interviewee 4)*

*“Well, I don’t know if this exists in general or Finland since we are in Finland. But is there a law that nonprofits have to report? [...] Is there a need? Because you know, when you are a company, you have to report all of these numbers, annually or quarterly; I wonder if there is anything that has ever existed or if it exists for nonprofits, some rule where we have to report so that you know uh, that we have this*

*data. So, if someone from another nonprofit wants to research, they can have this data. I don't think it exists, but that would be something to think about going forward for the future of nonprofits, like, you know, a reporting system for nonprofits, almost like a database with all this information.” (Interviewee 8)*

Once again, this group is identified only by a few of the respondents and not all, but as interesting points are made and the user group is validated by both types of nonprofits participating in the study, it is relevant to make the connection in the final framework.

To conclude this section, it is important to remember that from the list of users identified above; some will become intended users, meaning that they will be the ones for whom the evaluation should be conducted and the ones who should take part in the evaluation process for a better chance of a successful outcome and of concrete use of the evaluation. Other users will become unintended users and thus, according to the literature, should be omitted in the construction of the evaluation. Not all the users and uses present in the tool will happen simultaneously, yet having them in the final framework informs the nonprofits interested in using such framework of all the possibilities faced during the assessment phase of their program evaluation.

### **4.3 Strategic learning**

The place of strategic learning in program evaluation is given a strong position in this research because it has been proven to be the link between evaluation and strategy and a guiding line towards a concrete use of the findings (Williams, 2014). Surprisingly, the interviewees from all the organizations participating in this study had extensive thoughts about this subject. Initially, the conversation started around where and when the link should be established, meaning at which point the interviewees considered the results presented by the evaluator and the possibility that another evaluation could be needed. All the respondents stated that waiting and supporting an adjustment phase is necessary. A few persons were vague when discussing the concrete timeframe of evaluations; this fact was discussed with the concerned parties, and the reader may note that the cloudy

timeline is the result of the person not wanting to disclose information they were uncertain about; thus, indistinct timelines were discussed rather than precise ones.

*“First, we need to know what needs to be improved and what time frame.” (Interviewee 3).*

*“Maybe I would wait a bit before the next one. Like, uh, it depends on how much work there is to adapt from the evaluation. [...]. But, maybe after the results, so you have gotten them, you can process them and change some things.” (Interviewee 4).*

*“Maybe once a year [...], uh, it depends on the changes needed; if someone says that the whole system needs to change, then that’s a lot and a different story. But if it is something small [...], then that would be easier to change.” (Interviewee 5).*

However, most interviewees had a definite idea of the time frame needed for themselves or their organization to adapt to the recommendations and how regularly evaluations should be implemented. Concrete examples were shared during the interviews due to past experiences or personal knowledge. Alongside their answer on the place of strategic learning in program evaluations, some of the decision-makers provided insights on the factors that could be impacted by the learning phase, i.e., the place where the results are provided and reviewed by the organization.

*“... if it is the overall service, or if it is the people or the customer service, then the answer is not the same. But I think evaluations should be long, so there’s a line to follow; for example, in 5 years, we need to accomplish this or that. And then we can show that the goals are that, and the evaluation shows this, and we need that.” (Interviewee 1).*

*“So, it depends on the theme of the evaluation. Some should be yearly, every half a year, or something like that, but some [...] you should only do it once, and then the next one about this same theme is when the planning is already routine.” (Interviewee 2).*

*“I would wait for the results and then adjust before doing another. I wouldn’t wait too long because, I mean, the whole point of doing the evaluation is to spot the issue [...]. So, for example, if we evaluated Q1, we have the results a week into Q2, so we know that after March, in early April [...], we can know what adjustments to make, and then we try to put them in place in Q2-Q3. [...] it would take some months to implement the changes [...] and you would need to have the adjustments running for a least a quarter.” (Interviewee 6)*

*“If I do an evaluation and have done some work around it, I would not do another one soon after. I would try to adapt the work to the results of the first evaluation and see how it works. I would have to do an evaluation, maybe once a year, because there are many changes all the time, so we would have to do it regularly for it to be helpful for us. So, something regularly but with the time to adapt to the changes.” (Interviewee 7).*

*“Well, I think evaluations are [...] a way to know that you need to re-strategize, so I think it should be at least once a year [...]. Because maybe you need to re-strategize, but you don’t know until you have that data...” (Interviewee 8)*

Most respondents' conversations around strategic learning ended with the words above and insights. But with some of the interviewees, the discussion went deeper, and the interviewees shared thoughts about the need for evaluations in times of changes and uncertainty or wondered whether different evaluations could happen at the same time depending on the chance for those evaluations to focus on various aspects of the program at the same time. This research focuses on the singular place of strategic learning in general nonprofit program evaluations; additionally, only three interviewees disclosed either aspect voluntarily; as a result, such events are not represented in the final framework. Yet, the answers provide strong viewpoints and lines of research that could be investigated further, and they are worth mentioning here as the statements open the road to enrich further the question of strategic learning in nonprofit program evaluation.

*“...different things can be evaluated simultaneously, for instance, customer feedback, [our] service, etc. If several things are evaluated at the same time, then we can see if the evaluations agree or not with each other and see that okay, this is going well, or this needs to change, or do better, or if something is not even considered then we can see the priorities to take care of.” (Interviewee 1).*

*“But then, there could be other evaluations about how we do things, events, cooperation with partners, and [...] if I had time and money, then I would like to have a lot of different evaluations. [...] It is hard if there is only one evaluation; you always miss something.” (Interviewee 2).*

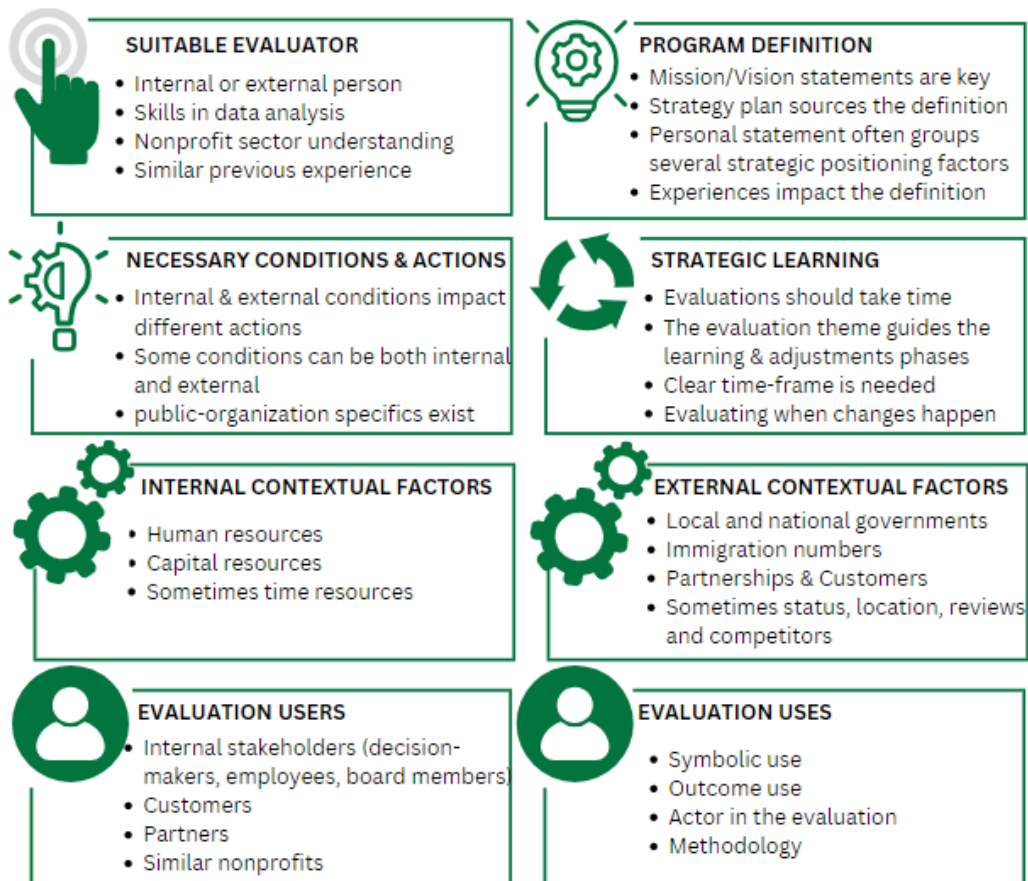
*“And then, another point is when there is a change, for example, we just re-strategized we did not know if it would work, and it sounded like a good idea, but in practice, it did not really work; so, we can have another evaluation when we launch it, just to see “ok does this idea work?” [...]. I would say evaluation should be launched before something changes, and when something changes [...], then make*

*it regular. Do you know how, like in a corporation, there is an annual review? Well, do the same thing here.” (Interviewee 8).*

The position granted to strategic learning is evident to all the respondents. They openly express and understand the necessity to adopt a learning curve carefully, followed by an adjustment phase in which changes are implemented to the strategy and the program specifics that follow the dissemination of evaluation results. Another evaluation can be considered only after the changes have been implemented for the participating nonprofits. This specific aspect of strategic learning is obvious in the answer as respondents mention the notion of routine several times, justifying that the adjustments and changes should become routine before another evaluation is considered.

#### **4.4 Summary of the results and revised framework**

The research aimed to analyze how nonprofit organizations can utilize program evaluation processes for strategic learning, and that was done through the combined use of the theory of change and the utilization-focused evaluation approach. Both perspectives have existing models, and when used separately, they provide different organizational learning. Yet, taken together, both methods complement each other, providing a line to follow for the evaluator and guidelines for the concrete use of findings. However, the study needed to gather answers from nonprofits to tailor the new framework, including both theories, for it to be used by and for nonprofits and their services. Altogether, the four case organizations participating in the study provided answers for the completion of the framework and recognized within their answers the eight different perspectives identified in the initial framework. Within these eight perspectives, several specific conditions tailored to nonprofits were pinpointed (Figure 12). Such identification is crucial for the evaluator to prepare and conduct the evaluation as it is the guideline to follow towards completion and use of the evaluation.



**Figure 12. Summary of program evaluation processes**

The findings recognized the need for evaluation in the nonprofit sector; clearly, some efforts are made by the organizations' boards to reach a certain level of knowledge through evaluation, but this often fails to meet the needs or remains surface-level. This was mentioned by a few to be linked to the evaluator used, but, overall, the question of the evaluator to conduct the research remains conflicted, some prefer the option of an internal evaluator while others spotted the risk of biases in choosing so. Regardless, the need for knowledge is there, and nonprofits wish to get many answers through the evaluation and results. Using the U-FE approach, the evaluator is guided through steps to identify and tailor the evaluation design to the primary users and uses. The respondents in the organizations have identified several possible users and uses of the evaluation.

Additionally, this study identified the connection between users and uses of evaluation. Identifying these two major aspects will ensure that the evaluation reaches the intended

target and goals. Using the ToC, the evaluator can quickly identify the link between the necessary conditions nonprofits must have to run the program and the actions conducted by the organization. The study’s participants have identified seven conditions required to run their program and actions smoothly; without those conditions, the program’s actions are at risk. In identifying those conditions during the evaluation assessment phase, the evaluator can already have a line of potential changes that need completion. The respondents' thorough answers identified links between conditions and actions.

Finally, by combining theoretical and practical approaches, the evaluator is led into different evaluation phases, from assessing program characteristics to designing, conducting, and facilitating the evaluation and results. Identifying the specifics of the ToC grants the evaluator knowledge of possible missing characteristics, while the UF-E provides in-depth guidelines on making a helpful evaluation for strategic learning. Together, a new framework (Figure 13) is created. The framework allows nonprofits to use program evaluation processes for strategic learning.

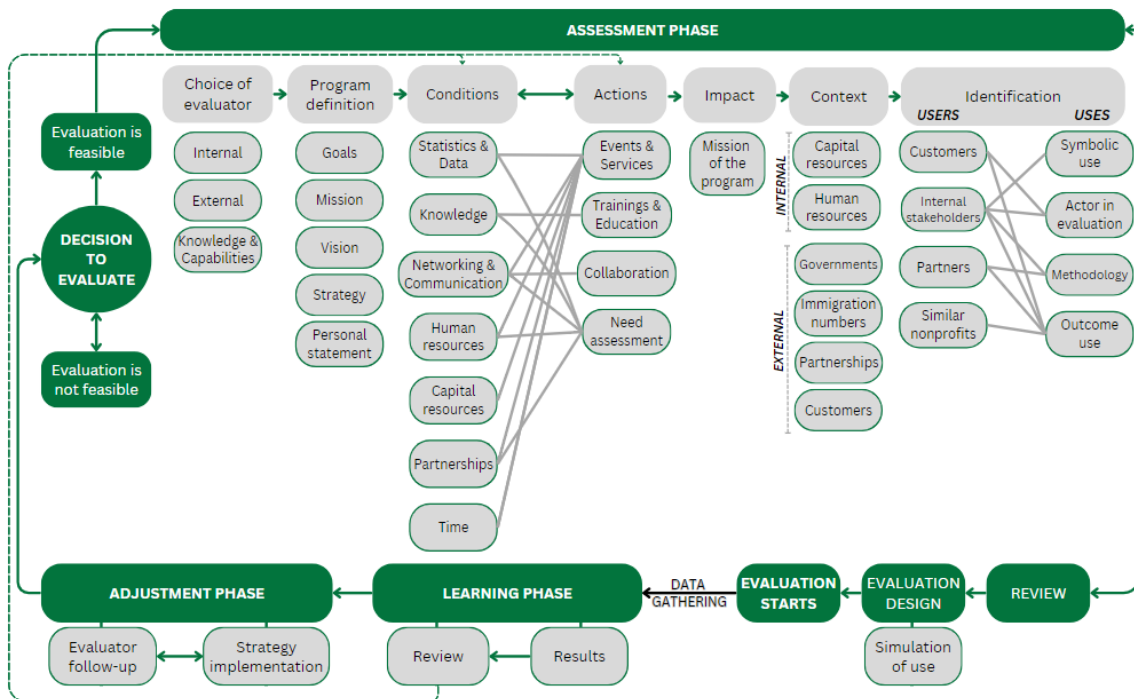


Figure 13. Revised conceptual framework

## 5 Discussion

### 5.1 Theoretical implications

As presented in the theoretical background, the concepts of strategic learning as a tool to reach extensive conclusions (Patton, 2008a) and using evaluations (Williams, 2014) exist in the name of program theories. However, further research demonstrates that linking evaluation results with strategic learning often remains surface-level (Williams, 2014) or is considered a linear process that forgets to include the concept of recurrence, sometimes used to critique the validity of evaluation results (Cook, 2000). This study extends the knowledge of program evaluation research by fulfilling the missed opportunity between using the evaluation and the link to strategic learning while providing an option for recurrence. The final framework utilizes program evaluation processes for strategic learning, including the timeline needed. The lack of guidelines regarding the timeframe to complete evaluation, as highlighted by Cook (2020), is shortly answered by the findings on strategic learning.

Additionally, the chosen field of study develops the learning in nonprofits. Indeed, they operate with specific differences from for-profit businesses (Benjamin et al., 2023; Carman, 2010) and with increasing demand for program theory (Rogers, 2007), yet the field sees an outstanding lack of it (Blair, 1998). So, by providing a new outlook, the research enriches the knowledge of nonprofit organizations.

The findings provided an overview of different theoretical aspects introduced in general program evaluation theory and highlighted which general theoretical factors are adaptable to the nonprofit field. Starting with the uses of program evaluation, it was clear that within the nonprofit sectors, not all the uses of program evaluation resonate. Only the outcome use, meaning the methodology and content of the evaluation (McDavid et al., 2013c), and the symbolic use of evaluation to prove the worthiness of the program (Falconer-Stout & Jones, 2021) were recognized as possible uses of the program evaluation in nonprofit organizations. Nevertheless, it is noted here that a distinction is made

between the outcome and content of the evaluation, described by McDavid et al. (2013) as one single use of the evaluation. This explains why the final framework distinguishes between outcome use and use of the methodology. Despite this, another use of program evaluation, unmentioned in theory, originated from the study and was named simply *actor in evaluation*. For nonprofit organizations, stakeholders are not the only party that should participate; some request customers to actively participate in the evaluation. According to Patton (2008a), this should also involve customers as the primary intended users of the evaluation, and the study confirms it.

Similarly, the users of program evaluations are not all plausible regarding the nonprofit sector. Two groups of users are identified in the literature: the users directly impacted by the evaluation (Benjamin et al., 2023; McDavid et al., 2013c) and those indirectly impacted by it (Weiss, 1998). The findings confirmed all the users from the first group, meaning the expected primary users, such as the leaders, decision-makers, funders, and people inside the organization (Benjamin et al., 2023; McDavid et al., 2013c). On the other hand, the extent of an evaluation conducted within the nonprofit sector appears to have limited reach in the second user group. Weiss (1998) expresses how far the evaluation may impact external users, but this research only validates the reach to similar nonprofits. Overall, the research results confirm the words of Benjamin et al. (2023) on the role of program evaluation towards decision-making purposes in nonprofit organizations and the statement of Patton (2005) on the power of identifying the correct users and uses of the evaluation.

Furthermore, the research study confirms the importance of having an adequate evaluator to the success of the evaluation (Cook, 2000; Williams, 2014) but contradicts Williams' (2014) mention of having an external person evaluating the most common option as the findings expressed that in the nonprofit sector, parity is more present in regards to the choice of evaluators.

Finally, the study settles theoretical arguments by creating a new framework combining the theoretical perspective with the ToC and a practical approach with the U-FE method. Factually, the final framework resolves the argument between Christie and Alkin (2003) and the reasoning that theory-based evaluation is more adapted to ensure utilization, and Scriven's (1998) and Stufflebeam's (2011) justifications that a single approach to evaluate risks endangering the validation and methodology of the evaluation. The final framework and the findings for the nonprofit organizations, thus consider thorough questions to understand the connections on how the program functions (McDavid et al., 2013e, 2013b) using two outlooks combined (Scriven, 1998; Stufflebeam, 2001).

## **5.2 Practical implications**

For nonprofit organizations, evaluating is a choice, not an obligation. Yet, the findings demonstrated that evaluations are needed, although rarely conducted in the studied sector. However, for public and volunteer organizations, choosing to conduct an evaluation is not easy, as there are many factors to consider before deciding to move forward. Indeed, depending on the choice of the evaluator, internal or external, having an evaluation may require more resources, either increased time and personnel or higher costs to consider in the given budget. Although the decision-makers could use the results, design, methodology, or any other aspects of the evaluation, in practice, before deciding to have an evaluation realized, there is an urgent need to weigh the pros of what the evaluation would bring against the cons of what the nonprofits must provide to complete the evaluation and facilitate the use of the findings.

Taking the road of evaluation would surely grant helpful knowledge for the board and the overall nonprofit. Such investigation could pinpoint whether the program is running as it should, if the necessary conditions for the nonprofit to run are present, and whether they are dedicated to the identified actions. Overall, different areas for change localized in one or more characteristics can be identified by evaluating the nonprofit programs. Additionally, evaluations can have unexpected consequences for nonprofits. The findings

confirmed that evaluations can be symbolically used, not towards program development per se, but as a means to an end. For example, the respondents explained that conducting an evaluation may help get more grants from governments, support their viewpoints in city councils, or help prove to the customers and members that the program is being developed. Finally, even without an external evaluator's help, the final framework's design allows nonprofits to spot critical elements strategically. Following the framework, nonprofit organizations may recognize if a necessary condition is missing or leading toward the linked actions. They may also adapt the framework to their programs by identifying the intended users and uses of the evaluation to facilitate the design toward concrete use.

Yet, one of the most important lessons gained from the research to answer the program evaluation question towards the strategic learning goal is that the evaluation process should be recurrent and implemented as a routine. Many published research and papers discuss how routine can improve task efficiency by enhancing both speed and reliability (Davis et al., 2009; Helfat and Peteraf, 2003; Nelson and Winter, 1982), allowing the program evaluation processes to become routine tasks within the public and volunteer organizations could enable decision-makers to reach more adequate decisions faster.

On a side note, when this research is being written, evaluations conducted by nonprofits are not reported, and every organization conducts its evaluations freely. However, a practical implication to be introduced here is that knowing that evaluation grants nonprofit boards crucial knowledge, reporting the evaluation and the practices used could further advance and help the sector of nonprofits. One interviewee mentioned, for example, the powerful impact that reporting evaluation and results in a database accessible by all nonprofits in the country would have for the nonprofit sector. Despite this, such practical implication requires more research.

### **5.3 Limitations**

Although this research has in-depth theoretical relations and practical answers, limitations remain. The suggested final conceptual model has not been tested in the targeted nonprofit sector; thus, a real-life test evaluating public or volunteer organizations would help identify which steps need rethinking and which could be validated. Furthermore, as the study was conducted recently after dramatic events that touched the European Union, i.e., the COVID-19 pandemic and the Ukrainian war, some of the respondents' answers might have been biased by those recent events. Similarly, the researchers' biases should be considered limitations to this study, as personal feelings or the lack of professional research experience may have impacted the findings.

Finally, several sub-factors may have impacted the study. As an inductive approach was taken for this study, the importance of context is undeniable (Saunders et al., 2007). This research is tailored to the contextual situation of Finland, with public and volunteer organizations dealing with the international population arriving or already living in the country. The reader is then advised that cultural aspects, respondents' position within the nonprofits, and previous work experience of the respondents may have influenced the results gained from this research.

### **5.4 Suggestions for future research**

The research aimed at answering how program evaluation processes for strategic learning could apply to nonprofit organizations and envisioned the answer through both perspectives of theory-based and utilization-focused evaluation combined into one framework. Nevertheless, this dissertation focused on the unique theoretical perspective of the theory of change in combination with a more practical approach. At the same time, other lines of theory-based evaluations exist, such as the realistic evaluations model, which is argued to provide unique views on complex social programs (McDavid et al., 2013b); interestingly enough, observing similar programs in a similar context through

another theoretical evaluation view would advance the knowledge in the field and grant a broader perspective on the topic. Whether combined with the practical approach of the U-FE or by itself, there is much to learn about program evaluation processes for the nonprofit sector. Additionally, studying similar practical cases but in another environment than hereby studied would enrich the program evaluation field and confirm or contradict the revised framework's use in other circumstances and conditions. Alternatively, as this dissertation followed a strict protocol of considering only the cases of public and volunteer organizations dealing with internationality and immigration in Finland, studying a different type of organization or sector in a similar environment would offer an additional viewpoint and an outlook on whether the revised framework presented in this research would apply to another type of organizations.

Finally, as this study was conducted in collaboration with the active levels and decision-makers of each program, it would be interesting to see the answers if similar questions were asked to the higher level, the ones granting the budgets, funds, or regulations to follow. As DuBow and Litzer (2018) highlighted, the program founders sometimes speak more regarding strategies, while others may relate more theoretical points, so the personal versus the practical perspectives could be an interesting road to study. Researching the differences in evaluation frameworks from the decision-maker level, studied in this dissertation research, and employees' or volunteers' levels might also deepen knowledge in the field of evaluation focused on nonprofits.

## 6 References

- Bellamy, E., & Evans, S. (2017). A study of outcomes determination strategy in non-profit health and welfare programs evaluation. *Evaluation Journal of Australasia*, 17(3), 23–31.
- Benjamin, L. M., Ebrahim, A., & Gugerty, M. K. (2023). Nonprofit Organizations and the Evaluation of Social Impact: A Research Program to Advance Theory and Practice. *Nonprofit and Voluntary Sector Quarterly*, 52(1\_suppl), 313S-352S. <https://doi.org/10.1177/08997640221123590>
- Blair, R. (1998). *Strategic planning for economic development: A suggested model for program evaluation*. 22(3), 313–348.
- Blamey, A., & Mackenzie, M. (2007). Theories of Change and Realistic Evaluation Peas in a Pod or Apples and Oranges? *Evaluation*, 13(4), 439–455.
- Bloom Ry. (2023). *Bloom ry*. Bloom Ry. <https://www.bloomtogether.fi>
- Breuer, E., De Silva, M. J., Fekadu, A., Luitel, N. P., Murhar, V., Nakku, J., Petersen, I., & Lund, C. (2014). Using workshops to develop theories of change in five low and middle income countries: Lessons from the programme for improving mental health care (PRIME). *International Journal of Mental Health Systems*, 8(1), 15. <https://doi.org/10.1186/1752-4458-8-15>
- Brousselle, A., & Champagne, F. (2010). Program theory evaluation: Logic analysis. *Evaluation and Program Planning*, 34(2011), 69–78.
- Bryson, J. (1988). Strategic Planning: Big wins and small wins. *Public Money & Management*, Autumn 1988, 11–15.

- Carman, J. G. (2010). The Accountability Movement: What's Wrong With This Theory of Change? *Nonprofit and Voluntary Sector Quarterly*, 39(2), 256–274. <https://doi.org/10.1177/0899764008330622>
- Chen, H.-T. (1996). A comprehensive typology for program evaluation. *Evaluation Practice*, 17(2), 121. [https://doi.org/10.1016/S0886-1633\(96\)90017-3](https://doi.org/10.1016/S0886-1633(96)90017-3)
- Christie, C. A., & Alkin, M. C. (2003). The User-Oriented Evaluator's Role in Formulating a Program Theory: Using a Theory-Driven Approach. *American Journal of Evaluation*.
- Coffman, J., & Beer, T. (2011). *Evaluation to support strategic learning: Principles and practices*. Center for Evaluation Innovation.
- Cook, T. D. (2000). The false choice between theory-based evaluation and experimentation. *New Directions for Evaluation*, 2000(87), 27–34. <https://doi.org/10.1002/ev.1179>
- Coryn, C. L. S., Noakes, L. A., Westine, C. D., & Schröter, D. C. (2011). A Systematic Review of Theory-Driven Evaluation Practice From 1990 to 2009. *American Journal of Evaluation*, 32(2), 199–226. <https://doi.org/10.1177/1098214010389321>
- DuBow, W. M., & Litzler, E. (2018). *The Development and Use of a Theory of Change to Align Programs and Evaluation in a Complex, National Initiative*. <https://doi.org/10.1177/1098214018778132>
- Dutton, J. E., & Duncan, R. B. (1987). The Influence of the Strategic Planning Process on Strategic Change. *Strategic Management Journal (1986-1998)*, 8(2), 103.

- Edmondson, A. C., & McManus, S. E. (2007). Methodological Fit in Management Field Research. *Academy of Management Review*, 32(4), 1155–1179. <https://doi.org/10.5465/AMR.2007.26586086>
- Eisenhardt, K. M. (1989). Building Theories from Case Study Research. *Academy of Management Review*, 14(4), 532–550.
- Eisenhardt, K. M., & Graebner, M. E. (2007). Theory Building from Cases: Opportunities and Challenges. *Academy of Management Journal*, 50(1), 25–32. <https://doi.org/10.5465/AMJ.2007.24160888>
- Eriksson, P., & Kovalainen, A. (2008). *Qualitative Methods in Business Research*. SAGE Publications Ltd. <https://doi.org/10.4135/9780857028044>
- EvalCommunity.com. (2024). *Understanding Formative Evaluation: Definition, Benefits, and Best Practices - EvalCommunity*. <https://www.evalcommunity.com/career-center/formative-evaluation/>
- Falconer-Stout, Z., & Jones, J. (2021). Utilization-Focused Evaluation Recommendations: Best Practices. *EnCompassing Evaluation*.
- Farquhar, J. D. (2012). Managing and Analysing Data. In *Case Study Research for Business*. SAGE Publications Ltd. <https://doi.org/10.4135/9781446287910>
- Franke, T. M., Christie, C. A., & Parra, M. T. (2002). Transforming a utilization focused evaluation (UFE) gone awry: A case of intended use by unintended users. *Studies in Educational Evaluation*, 29(1), 13–21. [https://doi.org/10.1016/S0191-491X\(03\)90002-2](https://doi.org/10.1016/S0191-491X(03)90002-2)
- Giancola, S. P. (2014). *Evaluation Matters: Getting the Information You Need from Your Evaluation*.

- Giancola, S. P. (2020a). Chapter 1 • Evaluation Matters. In *Program Evaluation Embedding Evaluation into Program Design and Development* (pp. 1–30). SAGE Publications, Inc.
- Giancola, S. P. (2020b). Chapter 2 • History of Evaluation. In *Program Evaluation Embedding Evaluation into Program Design and Development* (pp. 31–50). SAGE Publications, Inc.
- Gioia, D. A., Corley, K. G., & Hamilton, A. L. (2013). Seeking Qualitative Rigor in Inductive Research: Notes on the Gioia Methodology. *Organizational Research Methods*, 16(1), 15–31. <https://doi.org/10.1177/1094428112452151>
- Hastings, S. (1996). A strategy evaluation model for management. *Management Decision*, 34(1), 25–34. <https://doi.org/10.1108/00251749610106945>
- Hollister, R. (2003). *The Growth in After-School Programs and Their Impact*.
- International House Oulu. (2023). *International House Oulu*. Oulu.Com. <https://oulu.com/en/international-house-oulu/>
- Madaus, G. F., & Stufflebeam, D. L. (2000). Program Evaluation: A Historical Overview. In *Stufflebeam, D.L., Madaus, G.F., Kellaghan, T. (eds) Evaluation Models*. (Vol. 49, pp. 3–18).
- Maleka, S. (2014). *Strategy Management and strategic planning process*. 1.
- MAPorvoo. (2017). *Multicultural Association of Porvoo*. <https://www.maporvoo.fi/>
- McDavid, J. C., & Huse, I. (2012). Legislator uses of public performance reports: Findings from a five-year study. *American Journal of Evaluation*, 33(1), 7–25.
- McDavid, J. C., Huse, I., & Hawthorn, L. R. L. (2013a). Chapter 1. Key concepts and issues in program evaluation and performance measurement. In *Program evaluation*

- and performance measurement an introduction to practice* (2nd Edition, pp. 3–40). SAGE Publications, Inc.
- McDavid, J. C., Huse, I., & Hawthorn, L. R. L. (2013b). Chapter 2. Understanding and applying program logic models. In *Program evaluation and performance measurement an introduction to practice* (2nd Edition, pp. 45–88). SAGE Publications, Inc.
- McDavid, J. C., Huse, I., & Hawthorn, L. R. L. (2013c). Chapter 10. Using performance measurement for accountability and performance improvement. In *Program evaluation and performance measurement an introduction to practice* (2nd Edition, pp. 375–410). SAGE Publications, Inc.
- McDavid, J. C., Huse, I., & Hawthorn, L. R. L. (2013d). Chapter 11. Program evaluation and program management, joining theory and practice. In *Program evaluation and performance measurement an introduction to practice* (2nd Edition, pp. 411–437). SAGE Publications, Inc.
- McDavid, J. C., Huse, I., & Hawthorn, L. R. L. (2013e). *Program evaluation and performance measurement an introduction to practice* (2nd Edition). SAGE Publications, Inc.
- Mintzberg, H., & Lampel, J. (1999). Reflecting on the strategy process. *Sloan Management Review*, 21–30.
- Msila, V., & Setlhako, A. (2013). Evaluation of Programs: Reading Carol H. Weiss. *Universal Journal of Educational Research*, 1(4), 323–327.
- Newcomer, K., & Brass, C. T. (2016). Forging a Strategic and Comprehensive Approach to Evaluation Within Public and Nonprofit Organizations: Integrating Measurement and Analytics Within Evaluation. *American Journal of Evaluation*, 80–99.

- Nilsson, F., Petri, C.-J., & Westelius, A. (2016). Strategic management control in theory and practice. In *Strategic Management Control—With a focus on dialogue* (pp. 15–26).
- Patton, M. Q. (1994). Developmental evaluation. *Evaluation Practice*, 15(3), 311–319.
- Patton, M. Q. (1997). *Utilization-focused evaluation: The new century text*. (3rd edition). Thousand Oaks, CA: Sage Publications.
- Patton, M. Q. (2005). The challenges of making evaluation useful. *Assessment and Public Policies in Education (Original: Ensaio: Avaliação e Políticas Públicas Em Educação)*, 13(46), 67–78. <https://doi.org/10.1590/S0104-40362005000100005>
- Patton, M. Q. (2008a). Chapter 1. Evaluation Use: Both Challenge and Mandate. In *Utilization-Focused Evaluation: 4th edition* (4th Edition). SAGE Publications.
- Patton, M. Q. (2008b). *Utilization-Focused Evaluation: 4th edition* (4th edition). SAGE Publications.
- Patton, M. Q. (2013a). The roots of utilization-focused evaluation. In *Evaluation Roots: A Wider Perspective of Theorists' Views and Influences* (Second edition, pp. 292–303).
- Patton, M. Q. (2013b). Utilization-Focused Evaluation Checklist. *The Evaluation Center of Western Michigan University*.
- Patton, M. Q., & Campbell-Patton, C. E. (2021a). Chapter 1. The what, why, how, who, when and where of utilization-focused evaluation. In *Utilization-Focused Evaluation* (5th Edition). SAGE Publications.
- Patton, M. Q., & Campbell-Patton, C. E. (2021b). *Utilization-Focused Evaluation* (5th Edition). SAGE Publications.

- Patton, M. Q., & Campbell-Patton, C. E. (2023). *Our Story*. Utilization-Focused Evaluation. <https://www.utilization-focusedevaluation.org/our-story>
- Porter, M. E. (1996). What is strategy? *Harvard Business Review, Reprint 96608*, 61–78.
- Quinn, J. B. (1981). Formulating strategy one step at a time. *Journal of Business Strategy (Pre-1986)*, 1(42), 42–63.
- Ramirez, R., & Brodhead, D. (2013). *Utilization focused evaluation: A primer for evaluators*. Southbound.
- Ramirez, R., Galin, K., & Shephard, D. (2017). Utilization Focused Developmental Evaluation: Learning Through Practice. *Journal of MultiDisciplinary Evaluation*, 11(24), 37–53.
- Rogers, P. J. (2007). Theory-based evaluation: Reflections ten years on. *New Directions for Evaluation*, 2007(114), 63–67. <https://doi.org/10.1002/ev.225>
- Rossi, P. H., Lipsey, M. W., & Henry, G. T. (2017). *Evaluation: A systematic approach*. <https://catalyst.talis.com/uwasa/books/9781506307879>
- Rudov, L., McCormick-Rickett, I., Kingsmill, D., Ledford, C., & Carton, T. (2017). Evaluation recommendations for nonprofit social marketing campaigns: An example from the Louisiana Campaign for Tobacco-Free Living. *International Journal of Non-profit and Voluntary Sector Marketing*, 22(1), e1570. <https://doi.org/10.1002/nvsm.1570>
- Sassia, M., Jyrämä, A., & Pihlak, Ü. (2019). Using the Strategy Tripod to Understand Strategic Management in the “Evaluation-Friendly” Organizations of Cultural and Creative Industries. *Journal of Arts Management, Law & Society*, 324–346.

Saunders, M., Lewis, P., & Thornhill, A. (2007). *Research Methods for Business Students* (4th Edition). Pearson Education UK. <http://ebookcentral.proquest.com/lib/tritonia-ebooks/detail.action?docID=5139642>

Saunders, M., Lewis, P., & Thornhill, A. (2019). *Research Methods for Business Students* (8th edition). Pearson. [https://www.pearson.com/nl/en\\_NL/higher-education/subject-catalogue/business-and-management/Research-methods-for-business-students-8e-saunders.html](https://www.pearson.com/nl/en_NL/higher-education/subject-catalogue/business-and-management/Research-methods-for-business-students-8e-saunders.html)

Saunders, M., Lewis, P., & Thornhill, A. (2023). *Research Methods for Business Students* (9th Edition). Pearson International Content. <https://bookshelf.vitalsource.com/books/9781292402741>

Schwandt, T. (2015). *Evaluation Foundations Revisited: Cultivating a Life of the Mind for Practice*. Stanford University Press. <http://ebookcentral.proquest.com/lib/tritonia-ebooks/detail.action?docID=2038803>

Scriven, M. (1998). Minimalist Theory: The Least Theory That Practice Requires. *American Journal of Evaluation*.

Seinäjoen kaupunki. (2022). *City of Seinäjoki Strategy presentation 2022–2029* [Strategy presentation]. Seinäjoen Kaupunki. [https://www.seinajoki.fi/wp-content/uploads/2022/06/Seina%E2%95%A0ejoki\\_strategia2022\\_saavuttava\\_WEB\\_low.pdf](https://www.seinajoki.fi/wp-content/uploads/2022/06/Seina%E2%95%A0ejoki_strategia2022_saavuttettava_WEB_low.pdf)

Seinäjoen kaupunki. (2023a). *International Seinäjoki*. Seinäjoen Kaupunki. <https://www.seinajoki.fi/en/city-of-seinajoki/international-matters/>

- Seinäjoen kaupunki. (2023b). *More about MONI-Info service*. Seinäjoen kaupunki.  
<https://www.seinajoki.fi/hallinto/kansainvalinen-seinajoki/maahanmuuttajien-moni-info/moni-hanke/>
- Starbuck, W. H., Barnett, M. L., & Baumard, P. (2008). Payoffs and pitfalls of strategic learning. *Journal of Economic Behavior & Organization*, 66(1), 7–21.  
<https://doi.org/10.1016/j.jebo.2007.02.003>
- Stufflebeam, D. (2001). Evaluation Models. *New Directions for Evaluation*, 2001(89), 7–98. <https://doi.org/10.1002/ev.3>
- Thompson, J. D. (1967). *Organizations in Action: Social Science Bases of Administrative Theory*. New York NY: McGraw-Hill.
- Weiss, C. H. (1995). Nothing As Practical As Good Theory: Exploring Theory-Based Evaluation for Comprehensive Community Initiatives for Children and Families. In *New approaches to evaluating community initiatives* (pp. 65–92).
- Weiss, C. H. (1997). Theory-based evaluation: Past, present, and future. *New Directions for Evaluation*, 1997(76), 41–55. <https://doi.org/10.1002/ev.1086>
- Weiss, C. H. (1998). Have we learned anything new about the use of evaluation? *American Journal of Evaluation*, 19(1), 21–33.
- Whittington, R., Scholes, K., Angwin, D., Johnson, G., & Regner, P. (2023). *Exploring strategy. Text and cases* (13th ed.). Pearson Education Limited.
- Wholey, J. (1993). Evaluation and Performance. *Australian Accountant*, 28–33.
- Wholey, J. S., Hatry, H. P., & Newcomer, K. E. (2010). *Handbook of Practical Program Evaluation*. John Wiley & Sons.

- Williams, A. (2014). *Evaluation for Strategic Learning: Assessing Readiness and Results*. Center for evaluation innovation. <https://www.evaluationinnovation.org/publication/evaluation-for-strategic-learning-assessing-readiness-and-results/>
- Wiréhn, H., & Haag, S. (2020). *Applying systemic perspectives and lean to strengthen utilization-focused evaluation (EES16-0266)*.
- Wolf, C., & Floyd, S. W. (2017). Strategic Planning Research: Toward a Theory-Driven Agenda. *Journal of Management*, 43(6), 1754–1788. <https://doi.org/10.1177/0149206313478185>
- Yin, R. K. (2009). *Case study research: Design and methods*. (4th Edition). Sage Publications Inc.
- Zaveri, S., & Solomon, C. (2011). *Introduction to Utilization-Focused evaluation*. SLEVA, Colombo.

## Appendices

### Appendix 1. Interview Questions

#### BASIC INFORMATION

1. What is your role in the organization?
2. How long have you been working in this program?
3. Have you seen an evaluation conducted in your organization?
  - a. If yes, did you take part in it?

#### PROGRAM EVALUATION PROCESSES

4. What would you use to define the program:
  - a. Goals?
  - b. Mission statement?
  - c. Vision statement?
  - d. Strategy plan?
  - e. Personal statement?
  - f. Others?
5. What conditions are necessary for the program to run smoothly?
6. How does this link to the actions the program does?
7. Do you have an example?
8. If there were an evaluation, who would take part in its process?
9. Who would have access to the results?
10. What use would you make of the evaluation?
11. Which part of the evaluation would you use:
  - a. The results and conclusions?
  - b. The structure and methodology of the evaluation?
  - c. The sole idea of conducting an evaluation?
  - d. Others?
12. Which part of the evaluation would other people use?

- a. Results and conclusions?
- b. Generalities that apply to similar programs?
- c. The sole idea of conducting an evaluation?
- d. Others?

**GENERALITIES OF EVALUATION**

- 13. What external factors can impact the program?
- 14. What internal factors can impact the program?
- 15. Do you think anybody can conduct the evaluation?
- 16. Would you prefer someone internal or external to the organization?

**STRATEGIC LEARNING**

- 17. When would you decide that another evaluation is needed?
  - a. There would not be any other evaluation.
  - b. While the 1<sup>st</sup> evaluation is still in process.
  - c. When something changes, what is that "something"?
  - d. Right after getting the 1<sup>st</sup> evaluation results.
  - e. Right after adjusting the program based on the 1<sup>st</sup> evaluation conclusions.
  - f. Sometime after adjusting based on the 1<sup>st</sup> evaluation conclusions.
  - g. Other?

**PERSONAL DISCLOSURE**

- 18. Is there anything else you want to say about this topic?

## Appendix 2. Interviewees' summary

Interviewee number	Nonprofit organization type	Time in the organization*	Length of the interview
1	Public organization	Nine months	25m13s
2	Public organization	Ten months	56m08s
3	Public organization	3.5 years	24m14s
4	Public organization	Five months	25m07s
5	Public organization	Nine months	37m06s
6	Volunteer organization	2.5 years	34m22s
7	Volunteer organization	Five years	31m33s
8	Volunteer organization	Three years	25m34s

*\*The time in the organization is counted based on the interviews.*