Impact Measurement in Social Projects in Brazilian Communities: Bridging Everyday Evidence and Funder Expectations

Sarah Margulis

Dissertation written under the supervision of professor Fabricio Stocker

Dissertation submitted in partial fulfilment of requirements for the MSc in International Management - Entrepreneurship & Innovation, at Universidade Católica Portuguesa and for the MSc in International Management – Sustainability in Emerging Markets at Escola Brasileira de Administração Pública e de Empresas, at Fundação Getúlio Vargas, 04/01/2026.

Abstract

Title: Impact Measurement in Social Projects in Brazilian Communities: Bridging Everyday Evidence and Funder Expectations

Author: Sarah Margulis

Small community-based social projects in Rocinha and Vidigal (Rio de Janeiro, Brazil) must evidence impact to meet funder expectations embedded in calls for proposals (CFPs). This thesis examines how these projects define impact, document change, and navigate funder requirements. Using a proportionality perspective, it develops a triangle lens, purpose (whether evidence is decision-useful), feasibility (sustainable for low resource teams) and legibility (credible to external actors). The study uses a multiple-case design combining 13 interviews with project leaders, three specialist interviews, participant-shared routine artefacts, and documentary analysis of ten anonymized CFP packages. Findings show that projects rarely start from predefined indicators; they rely on embedded routines, such as attendance tracking and conversations with families. These indicators support coordination and learning but translate poorly into standardized tables. CFPs, by contrast, emphasize target-driven indicators and extensive documentation to support comparability. This often shifts compliance work onto already thin teams, which can lead to more reliance on intermediaries, selective compliance, or withdrawal from opportunities altogether. The study points to more proportionate funding and reporting requirement, including mixed-format proof of impact combining simple indicator sets with brief narratives of change. These adjustments can widen access to resources for social community projects. The thesis contributes with empirical insight into everyday evidence practices and introduces a purpose–feasibility–legibility lens to diagnose and improve fit between project routines and funder requirements.

Keywords: Social Impact, Social Project, Call for Proposal, Institutional Pressures, Funder Requirements, Monitoring & Evaluation

Resumo

Título: Mensuração de Impacto em Projetos Sociais em Comunidades Brasileiras: Conectando Evidências do Cotidiano e Expectativas de Financiadores

Autora: Sarah Margulis

Pequenos projetos sociais de base comunitária na Rocinha e no Vidigal (Rio de Janeiro) precisam demonstrar impacto para atender às expectativas de financiadores em editais e chamadas públicas. Esta dissertação investiga como esses projetos definem "impacto", registram mudanças no cotidiano e lidam com essas exigências. A partir de uma perspectiva de proporcionalidade, o estudo propõe uma lente em forma de triângulo: propósito (uso da evidência para decisões), viabilidade (sustentação por equipes com poucos recursos) e legibilidade (credibilidade reconhecida por atores externos). Com abordagem qualitativa de múltiplos casos, a pesquisa combina 13 entrevistas com lideranças de projetos, três com especialistas, artefatos de rotina e análise documental de dez pacotes de editais anonimizados. Os resultados mostram que os projetos raramente começam por indicadores pré-definidos; apoiam-se em rotinas da execução, como registros de presença e conversas com famílias. Esses indicadores favorecem coordenação e aprendizado, mas se traduzem mal em tabelas padronizadas. Em contraste, os editais privilegiam indicadores orientados a metas e documentação extensa, transferindo o trabalho de conformidade para equipes pequenas já sobrecarregadas e aumentando a dependência de intermediários, causando um cumprimento seletivo ou desistência de oportunidades. O estudo aponta para exigências mais proporcionais de financiamento e prestação de contas, incluindo formatos mistos que combinem indicadores simples com breves narrativas de mudança. Esses ajustes podem ampliar o acesso a recursos para projetos sociais de base comunitária. A dissertação apresenta a lente propósito–viabilidade–legibilidade para diagnosticar e aprimorar o ajuste entre rotinas dos projetos e exigências dos financiadores.

Palavras-chave: Impacto Social, Projeto Social, Editais, Pressões Institucionais, Exigências dos Financiadores, Monitoramento & Avaliação

Acknowledgements

Doing a double degree between FGV EBAPE and Católica Lisbon School of Business & Economics has been an unforgettable experience. Both institutions shaped how I think about management and the practical realities behind social impact.

I would like to thank my supervisor, professor Fabricio Stocker, for his guidance, encouragement, and feedback throughout the research process. I also thank him for his support beyond the dissertation, as a professor across various courses at FGV EBAPE.

The specialists who gave me valuable insights that I will carry with me in all my future social impact work. Getting to learn this much while doing data collection was a luxury.

I would also like to express my sincere gratitude to the project leaders who generously shared their time and experiences. I am thankful not only that they did an interview with me, but also showed me their projects and told me their histories. Many of these conversations stayed with me beyond what could be included in the thesis. I hope this dissertation does justice to your work.

Finally, I am also deeply thankful to my family, who have supported me throughout my life and encouraged me to pursue this master's degree. No matter where in the world I am, they are always there supporting me.

Table of Contents

List of Illustrations

List of Abbreviations and Acronyms

AI	Artificial Intelligence
CEP	Comitê de Ética em Pesquisa (Research Ethics Committee)
CFP(s)	Call(s) for Proposals
CLT	Consolidação das Leis do Trabalho (Brazilian labour code)
CNPJ	Cadastro Nacional da Pessoa Jurídica (Brazilian corporate taxpayer registry number)
COREQ	Consolidated Criteria for Reporting Qualitative Research
CSR	Corporate Social Responsibility
ESG	Environmental, Social, and Governance
FGV	Fundação Getulio Vargas
FGTS	Fundo de Garantia do Tempo de Serviço (Brazilian severance fund)
GIIN	Global Impact Investing Network
GRI	Global Reporting Initiative
IRIS+	Impact Reporting and Investment Standards (IRIS+)
LGPD	Lei Geral de Proteção de Dados Pessoais (Brazil's General Data Protection Law)
M&E	Monitoring and Evaluation
MSc	Master of Science
NGO(s)	Non-Governmental Organization(s)
RCT	Randomized Controlled Trial
SDG(s)	Sustainable Development Goal(s)
SIA	Social Impact Assessment
SROI	Social Return on Investment
SRQR	Standards for Reporting Qualitative Research
UN	United Nations
WBCSD	World Business Council for Sustainable Development

1. Introduction

In Rio de Janeiro, many small, community-based social projects work with short funding cycles, lean teams and basic data tools. Under these conditions, it is difficult to track outcomes over time, build credible counterfactuals or meet the demand for external evidence parameters (Silveira, 2020). At the same time, funders and intermediaries aggregate results across portfolios and expect to have clear indicators across diverse initiatives (GIIN, 2022). Between these two poles, thin local data and strong pressures for comparability, a recurrent tension emerges: what small teams can realistically document does not always match what external audiences expect. This thesis examines this tension in Brazilian community projects. It asks whether impact measurement is feasible in these settings and, if so, which forms of measurement are both realistic for projects and legible to funders.

The term “social project” is used here in a Brazilian sense: an organized, time-bound set of activities aimed at a defined public to address a concrete social problem (Stephanou et al., 2003). This label intentionally covers formal and informal efforts, with or without earned revenues, provided they are mission-driven and goal-directed. It is broader than “social enterprise” and narrower than social movements or campaigns, which may mobilize without the planning logic, objectives, milestones, means of verification, expected of projects (Stephanou et al., 2003; Monteiro, 2010).

In this thesis, “impact” refers to changes in how people live, organize, and exercise rights, effects on culture, participation, cohesion, health and well-being, and the lived environment (Vanclay, 2003). Consistent with SIA guidance, outcomes should be identified through participatory engagement with affected stakeholders, based on what they consider important; reporting should be transparent about what is included and excluded; and it should make the distribution of effects visible, who benefits and who bears costs (Vanclay et al., 2015).

Outputs alone rarely capture the kinds of change social projects aim for, especially when outcomes are non-market and multidimensional. Rather than reducing outcomes to a single score, this thesis uses two lenses that keep these dimensions visible. A capabilities lens focuses on the expansion achieved in what people can effectively do and be (Nussbaum, 2011). A relational lens focuses on relationships co-produced between people, such as trust, belonging and cooperation, that underpin community functioning

but are often under-represented in output-focused reporting (Donati, 2014). Taken together, these lenses point toward evidence strategies that combine a small set of capability- and relation-oriented indicators with brief narratives of change, instead of trying to compress change into a single headline number.

The broader institutional environment helps explain why some forms of evidence become more visible than others. Funders, regulators and intermediaries work with their own accountability needs and professional norms. Their language encourages projects to adopt familiar templates and indicator systems that travel across portfolios (DiMaggio & Powell, 1983). International frameworks, such as standards for sustainability reporting and impact-investing taxonomies, contribute to this shared language (GRI, 2021; GIIN, 2019).

Sustainability standards and impact-investing metric sets define what should be disclosed and how, making it easier to compare projects across a portfolio but also narrowing attention to what fits these templates (GRI, 2021; GIIN, 2019). Monetization frameworks add another layer, by expressing selected outcomes in monetary terms, so that they can be read alongside financial results (Nicholls et al., 2012; Serafeim et al., 2019). Rather than describing these tools in detail here, the thesis uses them later as reference points for judging what is asked of small projects and which expectations are realistic in their context.

For small, community-based projects, feasibility largely determines what can be credibly documented. In practical terms, this is about the time staff can spare for tracking, the stability of records and the ease of reaching participants, rather than about ideal research designs. Guidance for low-resource settings therefore recommends monitoring that is embedded in delivery and built around a few indicators that teams can realistically maintain (Silveira, 2020; Ebrahim & Rangan, 2014).

This study focuses on small, community-based projects in Brazil that work within tight capacity constraints and often in areas where public service provision and administrative data are uneven. The interest is not in finding a single best method, but in understanding when different ways of producing evidence fit or misfit these realities.

In this thesis, fit refers to alignment between three elements: (i) purpose: what evidence is for (accountability, learning, or decision support); (ii) feasibility: the time, data and

skills available to small teams; and (iii) legibility: refers to whether evidence is readable, usable and credible to the audiences it needs to serve, both internally (for decisions and learning) and externally (for selection, reporting, and accountability). This purpose–feasibility–legibility triangle structures the analysis and is summarized in Figure 1.

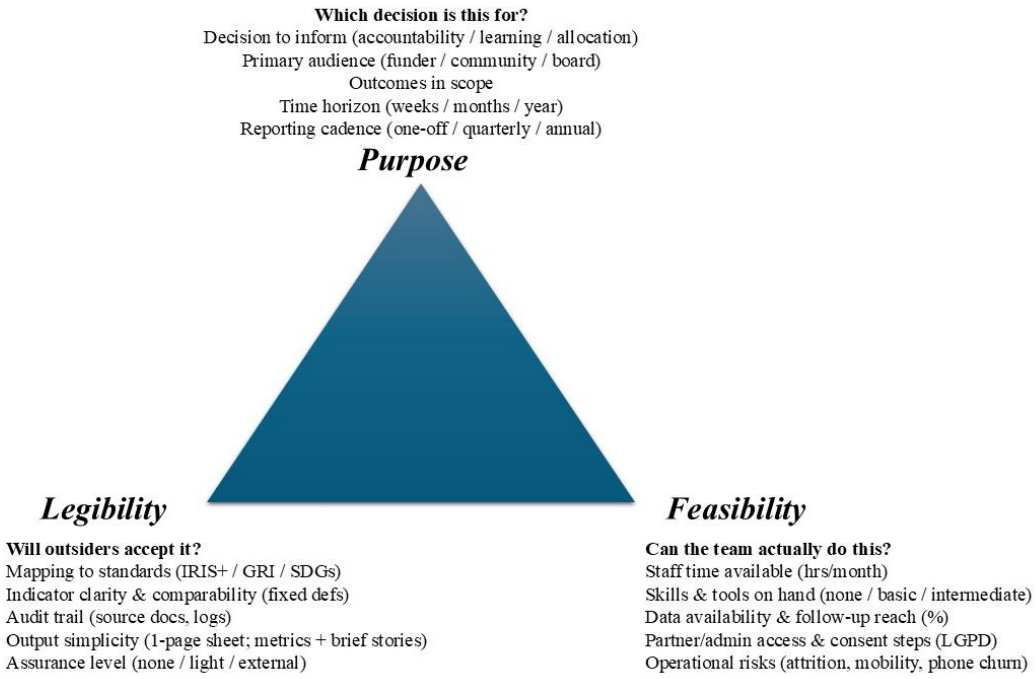

Figure 1. Purpose–feasibility–legibility fit check used in the thesis. Assessment focuses on the binding constraint, in other words the weakest criterion that limits the others (own figure).

The Brazilian setting also matters. The category of social project is widely used in Brazilian civil society and public social programs, and planning elements such as defined publics, objectives and indicators are common in guidance and training materials and widely used in practice (Stephanou et al., 2003; Prates Rodrigues, 2018; Monteiro, 2010). At the same time, many initiatives operate in communities with fragmented infrastructures and limited access to administrative data, which increases the value of stakeholder-scoped outcomes and qualitative evidence (Silveira, 2020; Vanclay et al., 2015). Throughout the thesis, community legitimacy is treated as a design requirement rather than an afterthought: indicators should make sense to residents first and only then be mapped to external taxonomies where this is helpful (Lacerda, 2014).

1.1 Research questions

The research is guided by three questions:

RQ1: How do small, community-based projects in Brazil define "impact" and show it with the resources they have in everyday practice?
RQ2: What do funders require in terms of indicators, documentation and monitoring, and how do project leaders experience these demands in practice?
RQ3: Based on this comparison, what proportionate adaptations or design principles could improve the fit between everyday evidence practices and formal measurement demands?

The thesis makes three contributions. Conceptually, it develops a purpose–feasibility–legibility lens that links institutional pressures with capability- and relational-oriented perspectives and brings funding rules into this picture as a concrete expression of legibility demands. Empirically, it offers comparative, multi-case evidence from Brazilian community-based projects, documenting how small teams define impact and navigate external reporting under real data and capacity constraints. Practically, it sketches proportionate adaptations or small "conversion recipes" that start from everyday evidence and suggest light ways of packaging it, together with a simple scoping rubric rather than a fixed template, so that funder asks can be aligned with on-the-ground feasibility while still remaining legible.

1.2 Objectives of the study

The main objective of this thesis is to understand how small, community-based social projects in Brazilian urban peripheries can produce evidence of impact that is feasible in their everyday realities and still legible to funders and other external audiences.

To pursue this objective, the study first seeks to reconstruct how selected projects themselves define "impact" and show their results with the resources they have. The focus is on how leaders talk about change, which outcomes they consider most important, and which routines and artefacts they already use in practice to follow participants and activities over time.

A second objective is to characterize the formal measurement and documentation demands that reach these projects through key funding instruments and support organizations. This involves examining how selected CFPs and related guidance describe

objectives, indicators, monitoring, reporting and prior experience, and how specialists in the field interpret these expectations.

Building on these two descriptions, the thesis then aims to analyze where everyday evidence practices and formal demands align or misalign when viewed through the purpose–feasibility–legibility lens introduced in the theoretical framework. Here the interest is less in compliance with any single method and more in how different actors negotiate what counts as "good enough" evidence under real constraints of time, data and capacity.

Finally, the study aims to gather design principles for impact measurement arrangements in small, community-based projects, drawing on cross-case patterns and specialist insights. Taken together, these objectives link the empirical chapters back to the central concern of the thesis: how questions about "measuring impact" in small projects are, in practice, questions about negotiating purpose, feasibility and legibility in unequal institutional settings.

1.3 Justification

This thesis is justified, first, by the practical importance of the projects it considers. Small, community-based initiatives in Brazilian urban peripheries often operate where public provision is uneven and families face cumulative forms of vulnerability. They provide everyday support in education, sports, culture and social assistance, yet usually work with short funding cycles, small teams and basic data tools. In this context, requirements attached to funding and partnership selection may influence which projects access resources and visibility. Understanding how these projects already show results, and how this compares to what funders formally ask for, is therefore relevant for designing support and accountability arrangements that do not systematically disadvantage smaller organizations or overlook forms of value that are harder to quantify.

The study is also justified on analytical grounds. Debates on social impact measurement and evaluation have largely been driven by examples of larger programs, social enterprises or impact funds, often in data-richer and more formalized settings. By focusing on small, community-based projects in Brazil and on the role of public and private CFPs in structuring evidence and reporting demands, the thesis brings together strands of literature on impact evaluation, social impact assessment and institutional pressures in a setting that remains under-represented in international discussions.

Finally, the research is justified by its feasibility. The design builds on access to project leaders, specialist informants and publicly available documentation, which makes it possible to compare everyday evidence practices and formal requirements within a clearly delimited set of cases and funding instruments. This combination of perspectives is intended to be useful both for practitioners who want fairer ways of showing results and for funders and intermediaries who are looking for more proportionate, context-aware approaches to impact evidence.

1.4 Research Structure

This thesis is divided into eight chapters, as listed below.

Chapter 1 introduces the research problem, objectives, and research questions, and outlines the overall approach.

Chapter 2 presents the research context, describing the ecosystem of small community projects, the role of funders and intermediaries, and why measurement is challenging in this setting.

Chapter 3 reviews the literature on impact measurement for small social projects in Brazil and synthesizes key tensions between everyday evidence practices and funder demands.

Chapter 4 develops the theoretical framework and introduces the purpose–feasibility–legibility lens used to analyze fit and misfit between practice and institutional requirements.

Chapter 5 explains the methodology, including the multiple-case design, data sources, sampling, analytical strategy, and ethical considerations.

Chapter 6 presents the empirical results on how projects define impact, how evidence is generated in routines, what CFPs require, and where the main gaps emerge.

Chapter 7 discusses the findings and implications for funders, intermediaries, and projects, linking back to the research questions and the analytical lens.

Chapter 8 concludes with the main contributions, reflections, and directions for future research.

2. Research context

2.1 The ecosystem of small social projects

In Brazil, the idea of a “social project” has become a common way of organizing work in civil society and in public–social partnerships. Manuals and training materials usually describe projects as a time-bound set of activities aimed at a defined public, based on a diagnosis and structured with objectives, expected results, indicators and means of verification (Stephanou et al., 2003). This planning language is used by NGOs, community associations, foundations and public programs and provides a shared template for how work should be described and justified (Prates Rodrigues, 2011).

At the same time, many projects that use this language operate with very lean structures. Studies of the Brazilian third sector point to a large number of small organizations that depend on volunteers, have fragile administrative routines and face constant pressure to secure resources (Prates Rodrigues, 2018). Monteiro (2010), analyzing small-grants programs, notes that projects are often run by a few key people who combine technical work, fundraising and administration and rely on basic tools to keep track of activities. In urban peripheries, community-based organizations frequently emerge around concrete local problems, such as school dropout, limited cultural opportunities or lack of spaces for children and young people, and they adapt their activities as situations change (Lacerda, 2014).

Evaluation and monitoring appear inside this ecosystem in uneven ways. Stephanou et al. (2003) present evaluation as a phase in the project cycle, typically at the end, when results are checked against objectives and indicators. Authors who observe projects from inside social policies and services describe a more continuous, if informal, process. Silveira (2020) argues that monitoring and evaluation happen in the everyday work of teams, when they define criteria, record information and revisit their practices. She also notes that, in many organizations, monitoring ends up reduced to filling in required forms and sending basic indicators to funders, with limited time for collective reflection (Silveira, 2020).

For community-based organizations, Lacerda (2014) shows that accountability is not only, or even primarily, about formal reports. Projects are also evaluated by residents through their presence in the territory, the consistency of activities and visible changes in relationships and opportunities. This “social accountability” coexists with more

bureaucratic demands addressed to public authorities and donors (Lacerda, 2014). Monteiro (2010) reports similar patterns in small-grant contexts: grantees often use simple before–after comparisons, attendance lists and feedback from participants and partners to judge whether a project is "working", even when they do not label these practices as evaluation.

2.2 Funders and intermediaries

The Brazilian third sector has expanded in parallel with corporate social investment, private foundations and organized philanthropy. Prates Rodrigues (2011) shows how corporate social projects have become a strategic tool for companies that seek to signal social responsibility while structuring their contributions through projects with explicit objectives and indicators. In her overview of third-sector organizations and social projects, she describes a field in which foundations, institutes and NGOs interact through partnerships, funding arrangements and networks such as GIFE, and in which expectations around planning and evaluation circulate through guidelines, training and CFPs (Prates Rodrigues, 2011). Expectations around planning and evaluation also circulate through practitioner-oriented guidance that revisits commonly used tools such as the logical framework "marco lógico" (Prates Rodrigues, 2014). In parallel, tax-incentive mechanisms channel some corporate support through regulated programs, which brings additional eligibility rules and documentation requirements for both donors and projects (Receita Federal do Brasil, n.d., 2025).

From the vantage point of donors and program managers, supporting many small projects at once creates a need for standardized information. Reporting templates often ask organizations to present indicators in comparable formats (Prates Rodrigues, 2018). This "professionalization" agenda is presented as a way to improve transparency and effectiveness, but it also introduces technical language and routines that may not have originated in the projects themselves (Prates Rodrigues, 2018). Monteiro (2010) describes how small-grants programs use simplified forms and short reports yet still require basic documentation and some form of results description, which can be demanding for groups with little administrative support.

Silveira (2020) highlights that, in public–social partnerships, monitoring and evaluation obligations are often expressed through standard forms, periodic reports and indicator tables. She observes that front-line teams frequently experience these requirements as

external: something that has to be completed to maintain agreements, rather than a set of tools fully integrated into internal management (Silveira, 2020). Lacerda (2014) shows that, for community-based organizations, complying with formal requirements sits alongside the need to maintain legitimacy vis-à-vis residents and local partners, which is assessed through different criteria.

In this environment, intermediaries play an important role. Prates Rodrigues (2018) points to the growth of organizations and consultants that support planning and evaluation for third-sector projects, often funded by donors who see this as capacity building. Their work includes helping teams clarify objectives, define a small set of indicators, design simple tools and compile information for reports. Silveira (2020) notes that external support can help insert moments of reflection and learning into routines that would otherwise be dominated by urgent service demands. She also warns that evaluation models imported without adaptation can overload teams or produce documents that travel well but remain only loosely connected to everyday practice (Silveira, 2020).

Taken together, these accounts depict a funding and support landscape in which small projects interact with larger institutions through project formats, CFPs and reporting templates. Expectations about evaluation and impact are articulated in these instruments, but the resources and support to meet them are unevenly distributed.

2.3 Why measurement is hard here

Against this backdrop, several authors underline why measuring results is particularly challenging for small, community-based projects. On the operational side, Silveira (2020) emphasizes that teams already work under heavy pressure to deliver services and handle administrative tasks. Time for systematic data collection, cleaning and analysis is scarce, staff turnover is frequent and information is often scattered across notebooks, forms and informal records (Silveira, 2020). Monteiro (2010) reports that small projects financed through small-grants programs frequently lack dedicated staff for monitoring and evaluation and must fit any additional work into already tight schedules.

Analytically, many of the conditions assumed in impact evaluation manuals are rarely present. Lacerda (2014) notes that community-based organizations in favelas tend to work with small groups, open enrolment, flexible activities and overlapping interventions. Under these conditions it is difficult to construct control groups, maintain stable protocols or follow participants over long periods in a way that would satisfy strict causal designs.

Prates Rodrigues (2018) argues that, in much of the Brazilian third sector, baseline data are missing, records are incomplete and projects change in response to emerging demands, which complicates attempts to apply experimental or quasi-experimental models in a straightforward way.

There are also representational and normative questions about what counts as a meaningful outcome. Vanclay (2003) defines social impacts in broad terms, including changes in how people live and work, and experience their environment. Stresses that affected groups should have a say in identifying which changes are most significant. In the Brazilian context, Lacerda (2014) shows that residents often judge projects by whether young people remain engaged in positive activities, whether relationships in the neighborhood improve and whether families feel supported. Such dimensions are not always captured by standard indicator lists. Nussbaum's (2011) capabilities approach and Donati's (2014) work on relational goods likewise highlight outcomes such as being able to concentrate, persist, trust and cooperate, which are difficult to compress into a small set of quantitative measures.

Finally, the literature suggests tensions between different arenas. For front-line teams and participants, what matters is whether activities are meaningful and whether they see concrete changes in routines and relationships (Lacerda, 2014; Silveira, 2020). For funders and program managers, evidence must also support selection, justification and control at a distance, often through comparable indicators and documents (Prates Rodrigues, 2018). When the same limited data are expected to serve all these purposes at once, the risk is that reporting drifts towards what fits templates rather than what best reflects lived change.

This thesis starts from these frictions. Small projects operate with short funding cycles, scarce data and layered accountability demands. The following chapters examine, for a specific set of projects in Rocinha and Vidigal and a sample of funding instruments, how everyday evidence practices and formal requirements meet in practice. It also studies how different actors negotiate what counts as "good enough" evidence in this setting.

3. Literature review: impact measurement for small social projects in Brazil

This chapter reviews how impact measurement and evaluation have been discussed in relation to small social projects, with particular attention to the Brazilian third sector. Building on the contextual overview in Chapter 2, the focus here is on the debates and tools that shape expectations about how these organizations should "show results". The chapter synthesizes empirical studies on how organizations generate evidence of results and how funders express their measurement demands. Together, these elements provide the basis for the purpose–feasibility–legibility lens that structures the empirical analysis.

3.1 Evaluation and impact measurement in the Brazilian third-sector debate

In practitioner guidance for social projects, evaluation is framed as something that should be planned and integrated throughout the project, rather than treated as a single activity at the end (Stephanou et al., 2003). Silveira (2020) similarly argues that evaluation and monitoring should be present across all stages of a project, including planning, implementation, and follow-up, and that this requires making criteria explicit rather than leaving them implicit. She also warns that evaluation can easily become a purely formal routine used mainly to justify action, which limits its potential to support learning and improvement.

In Brazilian debates, evaluation is commonly defined as a systematic and intentional process of collecting and analyzing qualitative and quantitative information to support learning, decision-making, and action. When the focus shifts to impact evaluation in a stricter sense, the central question becomes whether the project caused the observed changes. This requires counterfactual reasoning, as well as data, time, and technical expertise, conditions that are often not available to small organizations (Prates Rodrigues, 2018).

Guides for socio-environmental impact evaluation note that simple before–after tracking within the participant group can still generate useful management insights, even when it does not establish causality. Stronger causal claims require more complex designs, such as comparison groups or randomization, which also increase technical and financial demands (Insper Metricis, 2018). Prates Rodrigues (2024) argues that demanding strict impact evaluations analyzes from third-sector organizations is often disproportionate in Brazil. Rigorous impact estimation depends on comparison groups, stable data, and sophisticated modelling that can take years and significant resources to execute.

International reviews echo these concerns. Kah and Akenroye (2020) show that organizations choose methods not only for learning but also for legitimacy, while resource constraints remain a recurring barrier to adopting complex frameworks. Feor et al. (2023) further highlight ongoing tensions between standardization and local relevance, reinforcing the need for proportional approaches that align measurement ambition with feasible data collection and the nature of each intervention.

3.2 Families of impact measurement approaches and what they assume

The literature does not treat "impact measurement" as a single approach. Instead, it describes a field populated by several families of methods, each linked to different purposes and assumptions. Mapping these families is useful to understand what kinds of evidence different actors consider credible and which demands they place on organizations. In line with international reviews (Mishra, 2018; So & Staskevicius, 2015; Kah & Akenroye, 2020), this section organizes the discussion into four broad groups: monetization and economic-return approaches, indicator and reporting frameworks, theory-based and participatory approaches, and experimental and quasi-experimental designs.

3.2.1 Monetization and economic-return approaches

Monetization approaches aim to translate social and environmental outcomes into monetary values so they can be compared more directly with costs and investment decisions. This includes social cost–benefit analysis and Social Return on Investment (So & Staskevicius, 2015). A newer proposal is impact-weighted accounts, which seek to incorporate monetized social and environmental impacts into accounting statements alongside conventional financial performance (Serafeim et.al,. 2019). Building on this idea, the Impact-Weighted Accounts Framework proposes an Integrated Profit and Loss statement that sits alongside financial accounts. It aims to quantify and monetize positive and negative impacts across stakeholder groups (Impact Economy Foundation, 2024).

Guides to SROI describe a process that begins by defining scope and clarifying how activities are expected to lead to outcomes for stakeholders. They then identify material outcomes and use financial proxies to value those changes (Nicholls et al., 2012). The estimates are adjusted for what would have happened anyway (deadweight), for the contribution of other actors (attribution), and for changes over time (drop-off), before

calculating a ratio between the net present value of benefits and the value of the investment (New Economics Foundation, 2009).

In the Brazilian practitioner debate, Prates Rodrigues argues that monetization can be more plausible in settings where programs are large enough to justify specialist work and where the data needed for modelling can be assembled. She also warns that monetized ratios can appear more objective than they are, since results often rest on multiple hypotheses and choices that are not always visible to non-specialists (Prates Rodrigues, 2018). This concern is reinforced in her later writing, where she questions the routine use of SROI-style requirements for the third sector and highlights how easily a seemingly simple calculation can become a “black box” of assumptions (Prates Rodrigues, 2024).

Empirical studies of SROI applications point to similar trade-offs. In their analysis of social enterprises in health and social care, Millar and Hall show that SROI can support communication and internal reflection, but also involves practical barriers and ongoing resource demands that can be difficult to sustain without external support (Millar & Hall, 2013). In a sanitation case in Tanzania, Kwizela et.al demonstrate that SROI results shift substantially when assumptions about deadweight, attribution, and drop-off are varied, underlining how sensitive monetized estimates can be to modelling choices (Kwizela et al., 2018). The critique becomes sharper when monetization is treated as a generic expectation for small organizations. In this scenario, producing a ratio often depends on external expertise and layered assumptions rather than on robust local outcome data (Prates Rodrigues, 2018). Lazzarini et.al similarly caution that monetization can lose precision when key results are intangible or when estimates rely on assumptions that are not necessarily realistic, and they note that SROI in particular can become less precise when it attempts to monetize as much as possible (Lazzarini et al., 2021).

At the same time, Brazilian technical guidance for impact businesses frames economic valuation as one option within a wider toolkit and explicitly directs readers to dedicated material for monetization (Lazzarini et al., 2022). In health economics, the distinction between cost-effectiveness analysis and cost–benefit analysis also illustrates that monetization is only one way of structuring decisions: cost-effectiveness keeps outcomes in natural units, while cost–benefit requires assigning monetary values to effects to compare benefits and costs directly (Brent, 2023).
For small, community-based projects, the key question that follows from this literature is

less whether monetization is coherent in theory and more whether its data and modelling demands are proportionate to what teams can realistically produce.

3.2.2 Indicator and reporting frameworks

A second family of approaches focuses on constructing indicator systems and reporting frameworks that make performance visible in a structured and comparable way. In the impact investing field, this includes taxonomies of metrics, scorecards, and rating approaches that help investors benchmark organizations and aggregate results across portfolios (So & Staskevicius, 2015). Standardized metric catalogues such as IRIS+ are designed to support consistent definitions and comparability, including "core" metric sets that can be applied across investments (Global Impact Investing Network, 2019).

In corporate and public-sector contexts, similar dynamics appear in sustainability reporting, where standards specify what should be disclosed on material topics and provide topic-based requirements that organizations can use to structure their ESG reporting (Global Reporting Initiative, 2021). Guides that link reporting to the SDGs encourage organizations to select relevant indicators and report progress in ways that are legible to external stakeholders (GRI, UN Global Compact, & WBCSD, 2015).

In Brazil, planning-oriented materials often treat indicators as part of the basic structure of project design, drawing on tools such as the logical framework and indicator matrices that connect objectives, actions, indicators, and means of verification (Stephanou et al., 2003). Prates Rodrigues revisits the logical framework for third-sector projects and argues that it can clarify the theory of change and support the translation of abstract goals into concrete indicators, targets, and verification sources, even if it is not always applied during the planning stage (Prates Rodrigues, 2018).

Brazilian technical guidance for impact-oriented ventures also recommends starting from a clear intervention logic and then selecting a small, feasible set of metrics that reflect operational performance and key results, with transparency about assumptions and data limitations (Lazzarini et al., 2022). Some guides further suggest relating project metrics to SDG themes to position them within widely recognized development agendas (Insper Metricis, 2018).

The strength of indicator frameworks is that they generate regular data that can be used for internal management while also translating activities into a language legible to funders

and regulators (So & Staskevicius, 2015). At the same time, both Brazilian and international authors emphasize selectivity: indicators should reflect what is material and decision-relevant rather than expanding into exhaustive lists simply because they exist (Global Reporting Initiative, 2021). Stephanou et al. also note that the pressure to quantify can privilege what is easiest to verify, even though indicators are only signs of reality rather than reality itself (Stephanou et al., 2003).

For small community projects, this creates a familiar trade-off. Indicator frameworks can help translate everyday practices into a form that funders understand, but they can also increase reporting burdens and shift attention away from outcomes that matter locally yet are difficult to capture in standard templates (Prates Rodrigues, 2018). The literature therefore suggests that the number of indicators, the cost of producing them, and who participates in defining them are central questions for both feasibility and relevance (Lazzarini et al., 2022).

3.2.3 Theory-based and participatory approaches

Theory-based and participatory approaches start by making explicit how an intervention is understood to work, and they draw on the perspectives of staff, partners, and community members in developing and refining that account of change (Rogers, 2014). Most Significant Change operationalizes this by collecting stories of change and then using a structured selection process in which stakeholder panels discuss which stories are most significant and why (Davies & Dart, 2005). Outcome Harvesting similarly works with narrative-form evidence, but centers on outcome descriptions that are then substantiated with knowledgeable informants in order to build a credible account of what changed and how an intervention plausibly contributed (Wilson-Grau & Britt, 2013).

Rogers describes Theory of Change as explaining how activities are expected to produce a sequence of results that contribute to intended impacts. He emphasizes that these pathways rest on assumptions that enable or constrain change (Rogers, 2014). Vogel similarly argues that mapping a logical sequence is strengthened by critical thinking about context, stakeholders' roles, and differing assumptions about how and why change might occur (Vogel, 2012). When these pathways are clarified, they help specify what should be observed, which intermediate outcomes matter, and which contextual factors should be considered in data collection and analysis (Rogers, 2014).

From the perspective of social impact assessment, Vanclay argues that identifying social impacts should be grounded in what stakeholders consider important in a specific context, rather than treated as a generic checklist. He frames participation as actively involving affected people and building feedback mechanisms that allow communities to see that they have been heard. He also highlights distributional and ethical dimensions, noting that impacts can generate conflict when perceived as unfairly distributed and linking social outcomes to questions of rights and equity (Vanclay et al., 2015).

Brazilian literature echoes parts of this orientation by framing evaluation and planning as a dialog process, emphasizing the involvement of the people connected to the problem and the actions being taken (Prates Rodrigues, 2018). Silveira similarly argues that qualitative and quantitative information can complement each other, and that method choices should reflect what is feasible and useful for monitoring and evaluation purposes (Silveira, 2020). For small community-based projects, theory-based and participatory approaches can therefore provide a bridge between everyday knowledge and more structured reflection, but they still require time, facilitation, and a minimum level of documentation to sustain collective learning (Vogel, 2012).

3.2.4 Experimental and quasi-experimental designs

A fourth family of approaches is centered on experimental and quasi-experimental designs aimed at estimating causal impact. In the impact evaluation literature, randomized controlled trials and other methods are presented as ways to construct credible counterfactuals and isolate the effect of an intervention from other influences (Gertler et al., 2016; Mishra, 2018). These methods are particularly attractive when decisions involve large-scale expansion or significant reallocation of resources, and they are often promoted in development and social policy debates as “gold standards,” especially in the case of RCTs (So & Staskevicius, 2015).

At the same time, the literature is clear that these designs come with demanding assumptions and implementation requirements. RCTs rely on controlled procedures and are difficult to adjust once underway, while quasi-experimental strategies depend on the plausibility of the counterfactual and on conditions such as stable trends and comparability between groups (So & Staskevicius, 2015; Gertler et al., 2016). In observational settings where randomization is not feasible, tools such as propensity score methods are often used to strengthen causal inference, but they still require good data and

careful design choices (Rosenbaum & Rubin, 1983). Across these approaches, credible implementation often depends on data infrastructure and, in many cases, on sample sizes that are not always feasible or cost-effective for organizations operating at small scale (Mishra, 2018). Even when quasi-experimental designs are used, standard approaches such as difference-in-differences can give a false sense of precision if serial correlation is not handled properly. In these cases, conventional standard errors can be biased downward, which can make effects look more statistically significant than they really are (Bertrand et al., 2002).

Brazilian authors have engaged critically with this family in the context of third-sector projects. Prates Rodrigues (2018) notes that experimental impact evaluations are frequently treated as a "the golden standard" but that they demand technical capacity, robust data and resources that many organizations do not have, making them difficult to generalize as a default expectation. In analyzes focused on community-based organizations, Lacerda (2014) similarly emphasizes that projects cannot be assessed "in isolation" from the territory and its dynamics, which complicates attempts to cleanly attribute observed change to one intervention alone.

These observations do not imply that causal questions are unimportant, but they suggest that insisting on formal experimental designs as a general standard for "serious" evaluation may be impractical for many small projects. Organizations may still reflect on what might have happened without their intervention, compare experiences across groups and collect before–after information, but they will often need lighter designs and a more modest language of contribution rather than precise causal attribution (Gertler et al., 2016; So & Staskevicius, 2015).

3.3 A Brazilian agenda for proportionate evaluation

Against this background, some Brazilian authors argue explicitly for an evaluation agenda that is proportionate to the size, complexity and resources of projects, rather than anchored in a single methodological standard. In this view, methods should follow questions and context, not the other way around. Prates Rodrigues frames evaluation as a practical agenda that should be compatible with what Organizations can actually do, and warns against importing models that teams have neither the time nor the conditions to adapt to their realities (Prates Rodrigues, 2018). In a similar spirit, Brazilian guidance for impact businesses distinguishes between "basic" monitoring and progressively more

robust levels of additionality verification, suggesting that many initiatives should start with simpler designs and reserve more demanding approaches for selected questions where the added effort is justified (Lazzarini et al., 2022).

In her discussion of "impact evaluation: fantasy or reality?", Prates Rodrigues argues that treating strict causal impact evaluation as a universal benchmark can turn evaluation into an exercise in appearance rather than usefulness. She highlights that, when the conditions for robust designs are not in place, what gets produced may look technical while still being fragile, and she criticizes the tendency to equate "impact" with methodological status (Prates Rodrigues, 2021). In her later critique of routine SROI and monetized impact demands, she makes a related point from another angle: rigorous "impact evaluation" is demanding in time, data and design, and small Organizations often do not have the resources to meet these expectations without relying on external support and strong assumptions. In this sense, proportionality is not only about methodological fit but also about what evaluation requirements are effectively rewarded, given uneven organizational capacity (Prates Rodrigues, 2024).

Silveira approaches proportionality through the everyday function of evaluation. She argues that monitoring and evaluation should generate knowledge that helps professionals adjust practice over time, rather than being reduced to formalities that mainly serve external reporting routines (Silveira, 2020). From small-grants practice, Monteiro similarly stresses that small-grant approaches require simple procedures for project management and monitoring that are accessible to community-based Organizations, especially where administrative capacity is limited and bureaucratic requirements can become a barrier (Monteiro, 2010). Taken together, these contributions support an understanding of "good" evaluation as something that creates learning and credible accountability without demanding a level of technical infrastructure that grassroots teams do not have (Silveira, 2020).

These Brazilian debates also align with principles articulated in the international social impact assessment literature. Vanclay emphasizes that social impacts should be understood through the concerns and values of affected groups, and that participatory processes are central for identifying what counts as significant change and for whom (Vanclay, 2003). Later guidance similarly stresses scoping and context-sensitivity, noting that not all information is relevant in every situation and that early stages should clarify

the likely scale of effort required (Vanclay et al., 2015). International reviews of impact measurement echo this concern with fit and feasibility, showing that Organizations face persistent constraints around time, skills and resources and that tool choice is shaped by external pressure as well as internal purpose (Kah & Akenroye, 2020). This combination of perspectives reinforces the idea that proportionality is not the second-best option, but a principle for matching evidence demands to real conditions of practice (Feor et al., 2023).

3.4 Research synthesis: evidence practices and measurement demands

There is a growing body of work describing how organizations produce evidence of results and how measurement demands are demonstrated in funding and policy mechanisms. This section synthesizes findings most relevant to small, community-based projects in Brazil, focusing on three themes: everyday evidence practices, funder and funding-mechanism demands, and misalignments and coping strategies.

3.4.1 Evidence practices in small and community-based projects

Empirical studies of community-based organizations and NGOs in Brazil show that results are often documented through simple, everyday practices. Lacerda (2014) emphasizes that social impact must be read in relation to concrete changes in the territory and in residents' conditions of life. He also raises questions about accountability, including to whom organizations are accountable and whether scarce resources are being used in line with local demands. These forms of "everyday evidence" are closely tied to relational accountability: organizations are recognized and evaluated by residents through their presence, consistency, and perceived effects in the community (Lacerda, 2014).

Silveira (2020) notes that, in many social projects and social policies, monitoring relies on basic indicators such as participant numbers and activity frequency. Other indicators include immediate outputs, qualitative notes and team discussions. Formal evaluation reports, when they exist, often compile these elements in narrative form for funders or public authorities, but the underlying information is rarely systematized in a way that supports complex analysis because of time constraints, staff turnover, and limited information systems (Silveira, 2020).

Monteiro (2010), analyzing experiences with small-grant programs, describes grantees that assess whether a project is working through light reporting structures and informal

feedback from participants and partners. These observations may remain as brief notes or conversations, yet they inform everyday decisions about whether to continue, adjust, or close activities (Monteiro, 2010).

3.4.2 Measurement demands from funders and funding mechanisms

On the side of funders and policy instruments, measurement demands are often written into selection criteria and reporting routines for projects. In Brazil, Prates Rodrigues argues that as the third sector expanded, expectations for evaluation increased, even though third-sector projects are usually smaller and operate with fewer resources (Prates Rodrigues, 2018).

Brazilian technical guides aimed at socio-environmental and impact-oriented initiatives encourage organizations to define a small set of key metrics and clarify how those metrics relate to the intervention logic. They also put forward *additionality* as a central question and propose different "levels" of measurement rigor, while stressing that any approach should be transparent about its limitations and trade-offs (Lazzarini et al., 2022).

In parallel, international field-builders promote standardized metric sets to enable aggregation and comparison across investments. IRIS+ is explicitly framed as an effort to improve comparability across portfolios by organizing indicators and encouraging more consistent reporting (Global Impact Investing Network, 2019). Related impact-investing discussions similarly treat indicator sets and scorecards as a way to support portfolio-level decision-making and comparability across very different investments (So & Staskevicius, 2015).

In corporate and public reporting contexts, standardization appears through sustainability reporting frameworks that specify disclosures and associated indicators for ESG topics. These frameworks also emphasize that indicator choice should be driven by material topics and stakeholder relevance, rather than by mechanically applying long catalogues of metrics (Global Reporting Initiative, 2021). SDG-oriented guidance similarly encourages aligning measurement choices with SDG themes and targets as a way to communicate results within a shared sustainability language (GRI, UN Global Compact & WBCSD, 2015).

At the same time, Brazilian practitioner writing warns that these expectations can become disproportionate when treated as generic requirements for the third sector. Prates

Rodrigues argues that routinely demanding monetized impact studies such as SROI from smaller Organizations tends to privilege actors with technical support and stable data, rather than those with stronger local presence and relational work (Prates Rodrigues, 2024). In a related vein, Silveira notes that growing demands from financing agencies for "control instruments" around quality and impact can shape evaluation practices toward accountability pressures, even when Organizations have limited time and capacity (Silveira, 2020).

3.4.3 Misalignments and coping strategies

The juxtaposition of everyday evidence practices and formal measurement demands in the literature suggests areas of potential misalignment. Organizations that work mainly with informal and relational forms of evidence may struggle to translate their results into the indicator and framework language expected in CFPs and reports. When expectations for sophisticated evaluation designs or monetized metrics are added, the gap may widen further. Empirical accounts point to different ways organizations navigate this gap. One response is to treat evaluation mainly as a formal requirement, producing documentation to justify action rather than using it as a learning process. In contrast, learning-oriented guidance emphasizes evidence for reflection and course correction over time (Silveira, 2020; Hassnain et al., 2021). This can lead to a situation where reporting exists on paper, but it does not necessarily strengthen reflection on what is changing and why. A second response is to rely on specialized external support when measurement expectations exceed in-house capacity. Kah and Akenroye's review, although not specific to Brazil, aligns with this broader picture by emphasizing that legitimacy concerns are a central reason why Organizations engage with social impact measurement and by noting limited evidence that measurement tools are consistently integrated into everyday practice Kah & Akenroye (2020). This helps explain why Organizations may adopt the language of formal measurement while still relying on more pragmatic and situated ways of judging progress in day-to-day work.

Overall, the literature suggests that misalignments are not only technical problems of method choice but also reflect power and recognition: whose evidence counts, and which ways of knowing are treated as credible in funding and policy arenas.

3.5 Linking to the purpose–feasibility–legibility lens

The literature reviewed in this chapter points to three cross-cutting dimensions that shape impact measurement for small social projects: purpose, feasibility and legibility. Together, they highlight a recurring concern with proportionality, understood as matching methods to what information is needed, what is possible to collect and analyze, and which audiences the evidence is meant to address.

Across the reviewed work, purpose is not treated as singular. Authors describe measurement as serving multiple aims at once, including accountability, learning, decision-making and legitimacy, which can pull evaluation practice in different directions depending on context and audience (Ebrahim & Rangan, 2014). Feasibility is discussed as a persistent constraint, as methodological ideals confront limited time, unstable funding and uneven documentation, particularly for small initiatives operating with scarce data and limited staff capacity (Silveira, 2020). Legibility refers to how results are made understandable and credible to specific audiences in institutional environments shaped by multiple logics and expectations, where different actors rely on different forms of evidence and recognized categories (Greenwood et al., 2011).

These dimensions provide a concise way to organize the tensions observed in the literature between everyday evidence practices and more formalized measurement demands, and between locally meaningful accounts of change and standardized reporting formats (Global Impact Investing Network, 2019). Reeder and Colantonio note that “impact” itself remains contested and interpreted in multiple ways across fields, which further complicates efforts to produce evidence that is simultaneously credible across audiences (Reeder & Colantonio, 2013). Taken together, the literature suggests that questions of purpose, feasibility and legibility are not separate, but interact in shaping what types of evidence become possible and recognized in practice.

Building on this synthesis, the next chapter sets out how these three dimensions are used as an analytical lens in this study, before the empirical chapters examine how they play out in concrete cases and funding instruments.

4. Theoretical Framework

4.1 Framing the problem

This thesis examines the fit between funder expectations and the realities of small, community-based projects. As discussed above, project results are often made legible through standardized and comparable indicators, because this supports aggregation and comparison across portfolios (Global Impact Investing Network, 2019). At the same time, small projects often create context-specific forms of value that do not reduce neatly to a single metric, especially when change is relational, cumulative, and uneven across participants (Ebrahim & Rangan, 2014). In this thesis, fit means alignment among purpose, feasibility, and legibility.

As seen above, Chapter 3 mapped impact measurement families and their practical demands. Chapter 4 explains why particular approaches can dominate even when they are a poor match for small, community-based settings. Institutional theory helps explain how certain evidence forms become privileged within a field, because adopting widely recognized practices can increase legitimacy and survival prospects under conditions of uncertainty (DiMaggio & Powell, 1983). Critical CSR work complements this by showing how accountability and reporting discourses can function as strategies of legitimation, shaping what kinds of "responsible" action are recognized and rewarded (Banerjee, 2008).

Social-innovation perspectives add that change in social sectors is often embedded in services and co-created with users, which complicates linear input–output accounts of value creation (Mulgan, 2006). A capabilities perspective then keeps attention on plural outcomes in terms of what people are actually able to do and to be (Nussbaum, 2011). A relational perspective makes visible goods that are produced and "consumed" in the relationship itself, such as trust, recognition, and belonging, which standard dashboards often under-represent (Donati, 2014). Together, these perspectives help explain why everyday evidence in small projects looks the way it does, and how it fits in funder templates.

4.2 Institutional theory: how rules, norms, and field pressures define "valid" impact evidence

Institutional theory helps explain why certain forms of evidence become recognized as "valid" within specific fields. Rather than treating impact measurement as a purely

technical choice, this perspective emphasizes how practices are shaped by shared rules, norms, and expectations that Organizations adopt to secure legitimacy and resources (DiMaggio & Powell, 1983; Meyer & Rowan, 1977). In institutionalized environments, Organizations may conform to established formats because these are widely valued and legitimacy-enhancing, not necessarily because they best support local decision-making (Meyer & Rowan, 1977; Greenwood et al., 2011).

4.2.1 Institutional convergence and legitimacy

A central mechanism is isomorphism, through which Organizations gradually converge on similar structures and practices (DiMaggio & Powell, 1983). DiMaggio and Powell (1983) distinguish three sources of institutional isomorphism. Coercive isomorphism results from formal and informal pressures, including dependencies on powerful actors and expectations attached to resource flows. Mimetic isomorphism reflects imitation under uncertainty, when actors model themselves on Organizations perceived as legitimate or successful. Normative isomorphism stems from professionalization, including shared training, professional networks, and the diffusion of norms about appropriate practice (DiMaggio & Powell, 1983).

Applied to impact measurement, these mechanisms help explain why reporting templates and indicator-based approaches become widespread across grant-funded and accountability-oriented fields. When funders, regulators, or professional communities signal what "good reporting" looks like, adopting recognizable formats can become a rational response to field expectations, even when those formats only partially match operational realities (DiMaggio & Powell, 1983; Meyer & Rowan, 1977). Kah and Akenroye (2020) similarly highlight legitimacy as a recurring rationale for social impact measurement, which can orient tool choice toward what is accepted and defensible to external audiences. Taken together, this clarifies why "legible" evidence forms can diffuse quickly. They reduce uncertainty and help Organizations demonstrate conformity with prevailing accountability norms (DiMaggio & Powell, 1983).

4.2.2 Loose coupling and organizational responses

Meyer and Rowan (1977) extend this argument by showing that formal structures often operate as institutionalized myths, meaning culturally valued models of rational organization. Incorporating these myths can strengthen an organization's legitimacy, stability, and survival prospects. However, when institutional expectations conflict with

technical or practical requirements, organizations may protect day-to-day work by loosening the connection between formal structures and core activities (Meyer & Rowan, 1977). In this account, organizations may maintain ceremonial conformity while decoupling practices internally. Meyer and Rowan (1977) also stress that such arrangements can be sustained through a "logic of confidence and good faith," rather than through constant inspection and control. In impact measurement terms, this helps explain how projects can appear compliant through standardized reports while relying on more informal or adaptive information practices internally (Meyer & Rowan, 1977).

Institutional theory also highlights institutional complexity, where organizations confront incompatible prescriptions from multiple institutional logics (Greenwood et al., 2011). Greenwood et al. (2011) define institutional complexity as arising when organizations face competing demands grounded in different logics, each prescribing what is appropriate and how success is judged. Under such conditions, responses are not uniform. Organizations may manage tensions through selective framing, blending, or differentiation of structures and practices to meet different audiences' expectations. For impact evidence, this makes it plausible that organizations sustain parallel routines, one oriented to external accountability and another oriented to internal learning and adaptation (Greenwood et al., 2011).

4.2.3 Implications for feasibility and proportionality

These dynamics also complicate any assumption that measurement tools are neutral instruments. Reeder and Colantonio (2013) show that debates about impact assessment hinge on interpretive choices about what counts as impact, how attribution and contribution are constructed, and what constitutes credible evidence. They also highlight tensions between calls for high "rigor," for example experimental designs treated as gold standards, and the limits of such approaches for complex, context-dependent social change. From this perspective, the push toward standardization can privilege what is easiest to evidence and compare, while keeping aside outcomes that are relational or emergent (Reeder & Colantonio, 2013).

Feasibility concerns sharpen this point, especially for small or resource-constrained organizations. Ebrahim and Rangan (2014) argue that it is not feasible nor desirable for all organizations to develop metrics at all levels of a results chain. They emphasize alignment between measurement systems and project objectives, noting that measurement

extending beyond an organization's operational frame can distract from core work, and that broader, system-level impact is often better assessed at the funder level rather than pushed onto individual implementing organizations (Ebrahim & Rangan, 2014). In Brazilian practitioner guidance, Silveira (2020) similarly cautions that monitoring and evaluation should not be conducted merely to fulfil formalities, but should be planned so that it contributes to project development. Prates Rodrigues (2018) adds that even when data are collected, they often remain dispersed and fragmented, limiting the production of a consolidated understanding of project trajectories and results. Together, these sources support a proportionality argument. What is "reasonable" to evidence depends on purpose and capacity, not only on externally preferred formats (Ebrahim & Rangan, 2014; Silveira, 2020; Prates Rodrigues, 2018).

Institutional theory therefore contributes to this thesis by clarifying how field-level pressures shape what is recognized as credible impact evidence, and why standardized evidence forms can persist even when they are loosely coupled to practice (DiMaggio & Powell, 1983; Meyer & Rowan, 1977). It also provides language for analyzing how organizations navigate competing logics, particularly external accountability demands alongside internal learning needs, and why hybrid measurement routines can emerge as a pragmatic response (Greenwood et al., 2011). Within the purpose–feasibility–legibility lens developed in this thesis, institutional theory primarily strengthens the legibility dimension by explaining how legitimacy pressures stabilize particular reporting formats, and why these formats can conflict with feasibility constraints and learning-oriented purposes (DiMaggio & Powell, 1983; Greenwood et al., 2011; Ebrahim & Rangan, 2014).

4.3 CSR and impact-funding logics: why standardization and auditability dominate

4.3.1 Critical CSR: responsibility, legitimation, and the politics of proof

Critical CSR scholarship argues that measurement and reporting are not merely technical exercises, but part of how responsibility is defined and governed in contested arenas. In this thesis, this lens is used to explain why social impact measurement is often pulled toward auditable, standardized forms of evidence. What counts as "credible impact" can become closely tied to what can be documented, compared, and defended to external audiences. Banerjee (2008) frames CSR, and related corporate citizenship and sustainability discourses, as shaped by narrow business interests, with the effect of narrowing how external stakeholders' interests are recognized. In parallel, Barnett et al.

(2020) argue that CSR research has produced extensive knowledge about firm-level consequences while making limited progress in assessing whether CSR initiatives deliver the societal good they claim.

From this perspective, standardization becomes attractive because it produces portable and defensible forms of proof. Banerjee (2008) links CSR reporting to legitimation work, including cases where polished CSR reports function as greenwashing rather than transparent accounts of underlying realities. Barnett et al. (2020) add that what is studied and compared in CSR is shaped by the availability of large, public secondary datasets, which pulls attention toward what can be consistently coded across firms. Applied to social impact measurement, this highlights a practical selection effect. Indicators and claims that travel well across organizations and settings tend to dominate, while outcomes that are harder to observe, attribute, or standardize risk being marginalized in reporting routines. Together, these critiques suggest that "credible" responsibility can become associated with producing reviewable accounts, even when these accounts remain partial proxies for substantive change (Banerjee, 2008; Barnett et al., 2020).

These dynamics align with the institutional mechanisms discussed in Section 4.2, but the emphasis here is on why auditability becomes normatively attractive in CSR-linked accountability environments. Meyer and Rowan (1977) argue that organizations incorporate formal structures because they are culturally valued and legitimacy-enhancing, and that these structures can become loosely coupled to day-to-day practice. Greenwood et al. (2011) further show that in fields characterized by multiple, and sometimes conflicting, expectations, organizations may respond through compartmentalization or hybrids, which stabilizes outward-facing accountability routines even when internal learning routines remain different. Read through a critical CSR lens, this helps clarify how responsibility may be demonstrated through the production of credible accounts, even when changes in core practices are harder to evidence or remain contested (Banerjee, 2008; Meyer & Rowan, 1977; Greenwood et al., 2011). This framing sets the stage for the next subsection on impact-funding contexts, where comparability and auditability operate as governance priorities across portfolios and shape what kinds of impact evidence are demanded.

4.3.2 Impact-funding logics: auditability, benchmarking, and comparability pressures

Impact-funding and impact-investing contexts further intensify pressures toward standardization because funders seek comparable information across diverse organizations and geographies. Reeder and Colantonio (2013) describe persistent demand for standardized metrics and benchmarking alongside recurring tensions in practice, including trade-offs between rigor and flexibility, and between attribution and practicality. Systematic reviews similarly highlight that tool choices in social impact measurement are often shaped by external expectations for legitimacy and accountability, not only by internal learning needs (Kah & Akenroye, 2020). At the same time, Feor et al. (2023) note that there is no international consensus on social impact reporting, and caution that comparability should not come at the cost of completeness given varied stakeholder needs and the complexity of disclosures. Taken together, this suggests that auditability and benchmarking demands can rise even when shared standards remain unsettled, and even when important dimensions of impact resist commensuration (Feor et al., 2023; Reeder & Colantonio, 2013).

4.3.3 Counterpoint: proportionality and decision-useful evidence

A managerial counterpoint in this literature is the argument for proportionality and decision-usefulness. Ebrahim and Rangan (2014) argue that measurement should align with mission and strategy rather than escalating uniformly toward causal proof, and they stress that expectations often outpace the time and resources available to implementers. Vogel (2012) similarly cautions that performance-management tools can become too rigid to support adaptive decision-making in complex settings, which makes learning-oriented use distinct from compliance-oriented use. In Brazilian guidance on social project evaluation, monitoring is treated as a practical and continuous activity embedded in everyday project management, and there is no single “right” model that fits all initiatives (Stephanou et al., 2003). Practitioner-oriented Brazilian scholarship also points to feasibility constraints, including dispersed and fragmented information systems, which limit the consolidation and use of data for evaluation even where documentation exists (Prates Rodrigues, 2018).

Taken together, CSR and impact-funding logics help explain why auditability and standardization often dominate evidence design, because they support reputational management and portfolio governance by producing comparable and reviewable accounts (Barnett et al., 2020; Reeder & Colantonio, 2013). At the same time, the critical CSR lens

warns that these “legible” formats can privilege what is easy to report over what is substantively significant for affected groups (Banerjee, 2008). This tension motivates the purpose–feasibility–legibility framing used in this thesis. Legitimacy-driven demands for standardized evidence can rise independently of feasibility, and can exclude learning-oriented documentation even when learning would be more decision-relevant for implementers (Ebrahim & Rangan, 2014; Vogel, 2012).

4.4 Social innovation, capability, and relational lenses: why grassroots value is hard to quantify and how to evidence it

4.4.1 Social innovation and emergent pathways

Work on social innovation treats change as something produced in practice, through ideas that “work” in the sense that they address social needs and generate improvements, rather than through linear execution of predefined plans. A key implication is that innovation is often a discovery process. Ideas are reshaped as they encounter reality, and progress emerges through iteration, testing, and feedback loops (Mulgan, 2006). This makes innovation difficult to capture through a single input–output chain, because the pathway itself is part of what is learned and negotiated over time (Howaldt & Schwarz, 2010). Theory-oriented approaches help make this evolving pathway discussable by foregrounding assumptions and contextual conditions that enable or constrain change, rather than treating causal links as self-evident (Rogers, 2014). Social innovation scholarship also emphasizes that what counts as “successful” depends on uptake within specific contexts and target groups, which limits the usefulness of uniform success criteria across settings (Howaldt & Schwarz, 2010).

4.4.2 Capability and relational value

The capability approach reframes “impact” as the expansion of real opportunities, meaning what people are effectively able to do and to be, rather than only the observation of achieved outcomes. This matters for evidence because well-being is not a single dimension. Reducing lived change to a commensurable scale can obscure differences that are ethically and practically significant (Nussbaum, 2011). A relational lens complements this by emphasizing forms of value generated through interaction, such as trust, recognition, belonging, and solidarity, where the “good” is jointly produced and experienced. Because these relational goods depend on time, reciprocity, and sustained involvement, they are often under-represented in dashboards that privilege easily

countable units (Donati, 2014). Social Impact Assessment (SIA) provides a useful bridge by defining social impacts broadly as changes in how people live, work, relate, and organize, which legitimizes attention to shifts that may not be reducible to simple proxy indicators (Vanclay, 2003).

4.4.3 Evidencing change through proportional, purpose-sensitive designs

These lenses do not imply that indicators are irrelevant, but they point toward evidence practices that are proportionate and attentive to mechanisms of change, as well as to what can realistically be observed and documented in a given setting. Brazilian planning-oriented guidance frames indicators as "signals" that point toward reality rather than substituting for it, and notes how verification pressures can privilege what is easiest to quantify and check, even when that emphasis narrows what is seen (Stephanou et al., 2003). Measurement design can therefore start from purpose. Evidence needs differ depending on whether the aim is estimating, planning, monitoring, or evaluation, and the same system rarely serves all objectives equally well (So & Staskevicius, 2015). Theory of Change supports this alignment by clarifying a sequence of results and making underpinning assumptions explicit, which guides what should be observed and how it should be interpreted (Rogers, 2014). Participation-based traditions in SIA reinforce this orientation by involving affected groups in identifying relevant issues and interpreting consequences, which helps keep evidence anchored in what stakeholders consider salient in context (Vanclay, 2003; Vanclay et al., 2015).

In low-resource settings, feasibility is not a secondary concern but part of methodological validity. Brazilian practitioner guidance emphasizes that evaluation demands time and energy and should not be treated as a formality; monitoring is more defensible when it functions as an ongoing learning process that supports course correction (Silveira, 2020). Where data systems are dispersed or fragmented, even reasonable reporting expectations can become difficult to meet consistently, which should shape what counts as a credible design (Prates Rodrigues, 2018). Prates Rodrigues further cautions that impact evaluation can drift into "fantasy or fiction" when methods are disproportionate to available data or capacity, strengthening the case for balancing purpose, feasibility, and legibility rather than escalating method ambition by default (Prates Rodrigues, 2021).

4.5 Analytical model: institutional demands, project realities, measurement choices and consequences

This section consolidates the theoretical framework into the analytical model used in this thesis. The model treats social impact measurement as shaped by the interaction between, first, institutional demands that define what counts as credible and reviewable evidence and, second, project realities that shape what can be meaningfully observed, recorded, and interpreted in practice. This interaction influences measurement choices, including what is measured, how it is represented, and for whom, with consequences for accountability, learning, and which forms of value become visible (DiMaggio & Powell, 1983; Meyer & Rowan, 1977; Greenwood et al., 2011).

4.5.1 Institutional pressures and infrastructures of legibility

Institutional theory suggests that organizations face pressures to adopt recognizable and defensible reporting practices, which can contribute to convergence in measurement formats (DiMaggio & Powell, 1983). Standards and intermediary infrastructures reinforce this by codifying what "good reporting" looks like and by privileging characteristics such as comparability and verifiability (Global Reporting Initiative, 2021). In impact investing, metric catalogues and core sets similarly aim to standardize reported outcomes and support comparison across diverse investments and contexts (Global Impact Investing Network, 2019). At the same time, synthesis work cautions that standardization does not remove underlying complexity. There is no international consensus on social impact reporting, and comparability can conflict with completeness when stakeholder needs and disclosure contexts vary (Feor et al., 2023).

These dynamics are reinforced by the institutional pattern that formal structures can be adopted because they are legitimacy-enhancing, even when they are loosely coupled to day-to-day practice (Meyer & Rowan, 1977). Where accountability demands are strong, organizations may stabilize outward-facing routines that are legible to external audiences, while internal learning practices remain different or more informal (Greenwood et al., 2011).

4.5.2 Why grassroots value resists standardization

The model also draws on lenses that explain why many grassroots outcomes are difficult to compress into stable, comparable indicators. Social innovation accounts emphasize that change processes are iterative and adaptive, shaped by experimentation, feedback, and diffusion rather than linear execution of predetermined plans (Mulgan, 2006; Howaldt &

Schwarz, 2010). The capability approach frames social value as expansion of real opportunities, meaning what people are able to do and to be, and cautions that well-being is plural and not easily reduced to a single commensurable scale without losing ethically and practically significant distinctions (Nussbaum, 2011). A relational perspective highlights that some central outcomes are produced through interaction, for example trust, belonging, and recognition, which makes them difficult to represent through dashboards that privilege easily countable units (Donati, 2014).

Social Impact Assessment provides a bridge by defining social impacts broadly as changes in how people live, work, relate, and organize, and by emphasizing attention to context, distributional effects, and stakeholder participation in identifying and interpreting significance (Vanclay, 2003; Vanclay et al., 2015). Taken together, these perspectives support the thesis' premise that what is substantively meaningful to affected groups may be only partially captured by the standardized forms of proof favored by external audiences.

4.5.3 Operationalizing the triangle

To analyze these tensions without assuming a single "best" method, the thesis uses a purpose–feasibility–legibility triangle.

Purpose refers to what the evidence is meant to support, for example accountability, learning, or resource allocation. Evidence designs are expected to differ depending on the decisions they are meant to inform (Ebrahim & Rangan, 2014; So & Staskevicius, 2015).

Feasibility refers to what can be credibly collected and analyzed given resources, data infrastructure, and skills. Brazilian practitioner-oriented sources emphasize that evaluation and monitoring require time and planning and should not be treated as formalities. Feasibility therefore becomes part of what makes evidence credible in practice (Silveira, 2020; Prates Rodrigues, 2018). Prates Rodrigues (2021) further warns that evaluation can drift into "fantasy" when methods and claims exceed what data, context, and capacity can sustain.

Legibility refers to what forms of proof are recognized as credible by salient audiences. Standardized principles and metric infrastructures can strengthen legibility by producing comparable and reviewable accounts (Global Reporting Initiative, 2021; Global Impact

Investing Network, 2019), but these may privilege what is easiest to quantify and verify (Stephanou et al., 2003).

This framing implies that method choice should be proportional and contingent. Theory-based approaches can strengthen purpose and feasibility by clarifying assumptions and intermediate outcomes, helping initiatives decide what is realistic to observe and how to interpret change over time (Rogers, 2014; Vogel, 2012). Indicator frameworks can strengthen legibility, but they require explicit attention to assumptions, limitations, and uncertainty if they are to remain credible rather than merely performative (Global Reporting Initiative, 2021). Where causal attribution is necessary and feasible, impact evaluation designs can be appropriate, but they require conditions and data that are not always present in small initiatives (Gertler et al., 2016).

In sum, the analytical model guides how the thesis interprets measurement practice. It examines how actors negotiate evidence demands and make trade-offs across purpose, feasibility, and legibility, and how these trade-offs shape what becomes visible as "impact" and what remains outside standardized accounts (Ebrahim & Rangan, 2014; Meyer & Rowan, 1977; Greenwood et al., 2011).

5. Methods

5.1 Research design and approach

This study adopts a qualitative, multiple-case design to examine how small, community-based social projects in Brazil define and evidence impact, how they experience external reporting demands, and how specific tools fit day-to-day realities. A case study approach is well suited when the aim is to understand a complex issue in depth, in its real-life context, especially when the boundaries between the phenomenon and its context are not clearly evident (Crowe et al., 2011). In this thesis, the phenomenon is how small, community-based projects produce and use impact evidence in everyday practice, and how this intersects with external reporting demands.

The design integrates multiple sources of evidence for each case, so practices and claims are not read through a single lens (Baxter & Jack, 2008). For each case, the study draws on more than one source of material, so the analysis does not rely on interviews alone. Combining interviews with documents and routine artefacts helps build a fuller picture

of what is happening and allows claims to be checked across sources (Baxter & Jack, 2008; Crowe et al., 2011).

Write-up follows the SRQR reporting guidance so that key methodological choices and analytic steps are described clearly and readers can assess the quality of the study (O'Brien et al., 2014). For the interview components, the study also draws on relevant COREQ items to ensure that important details about the interview process and researcher role are reported transparently (Tong et al., 2007). The analysis is interpretive and practice-oriented: it focuses on participants' meanings while aiming to produce insights that are useful for project teams and their supporters. The empirical material combines three strands: (i) thirteen semi-structured interviews with leaders of small community-based projects, (ii) three specialist interviews with practitioners and researchers working on social impact evaluation and support of social projects, and (iii) a structured documentary analysis of anonymized Brazilian Call for Proposals (CFPs) and their annexes. In parallel with interviews, the CFPs analysis makes funder expectations, eligibility, documentation, and monitoring and evaluation (M&E) requirements comparable across sponsor types and provides a benchmark against which project evidence and reporting are compared.

5.2 Setting, unit of analysis, and case boundaries

The setting is Brazil's third sector, with fieldwork conducted in the Rio de Janeiro communities of Rocinha and Vidigal. The unit of analysis is one social project, meaning one project corresponds to one case. This is consistent with qualitative case study guidance that defines the "case" as the bounded phenomenon under study and, in effect, the unit of analysis (Baxter & Jack, 2008). Case boundaries were defined to keep the phenomenon and scope manageable. They combine place (Rocinha and Vidigal), organizational focus (small community-based initiatives), and time, expressed as a minimum operating period (Crowe et al., 2011). Projects were eligible if they had been operating for at least twelve months at the time of inclusion, although most had been active for longer. Interviews were conducted with project leaders or coordinators who were currently responsible for the projects. In total, thirteen project-leader interviews were completed.

Specialist interviews form a complementary strand. Three specialists working with evaluation of social impact, support, or research on social projects in Brazil were

interviewed to provide an external vantage point on what counts as “good enough” evidence for small projects, which tools tend to be realistic or unrealistic in low-resource settings, and how broader funding arrangements shape visibility.

In parallel, the study assembled a contemporaneous corpus of CFPs, including ten public, corporate, and foundation CFPs available to the public. The materials include applicant instructions and guides, selection and scoring descriptions (including points systems), and application and question sets. These documents are not treated as cases. Instead, they are used as contextual rules and benchmarks against which project evidence and reporting are compared. To protect anonymity, sponsors and programs are not named, and CFPs are referred to generically, for example “a state socioenvironmental call” or “a corporate sports program”. A summary of the documents is provided in appendix A.

5.3 Sampling strategy and sample size rationale

Sampling was purposeful. Recruitment was facilitated by a community liaison in Rocinha and Vidigal, with limited snowballing from initial contacts. Participants were selected because they had direct, relevant experience with the phenomenon under study and could speak to it in depth, rather than to produce a statistically representative sample (Palinkas et al., 2015). No fixed size thresholds were set in advance. Approximate team size and volunteer share, as well as place-embedded delivery, were recorded descriptively for each case.

The initial plan was to reach proximally twenty interviews. In practice, thirteen project leaders and three specialists were interviewed. Sample size adequacy was assessed through the lens of information power rather than a fixed target, treating sample size as contingent on the study aim, the specificity of the sample, and the richness of the dialogue, in relation to the analysis strategy (Malterud et al., 2016). During fieldwork, decisions about whether to continue recruitment were informed by periodic checks of whether additional interviews were still adding substantively new issues or insights relevant to the research questions. When later interviews largely repeated what had already been captured, further data collection was judged unlikely to add meaningful dimensions (Hennink & Kaiser, 2022).

In parallel, a purposive documentary sample of CFPs was assembled for contrastive description. Without naming programs, selection prioritized variation by sponsor type (public, corporate, foundation), sector, and recency, focusing on documents in force

during fieldwork. The CFPs were included in the analysis only if their eligibility description explicitly mentioned “favela”, “Rocinha” and/or “Vidigal” as target locations. This criterion helped ensure that they were intended for projects comparable to those in this study. Inclusion required an accessible rule document, such as an Edital (CFPs) or Chamada Pública (public call), Regulamento (regulations), Guia do Proponente (Guide for Applicants), and anexos (annexes). Press releases without rulebooks and superseded drafts were excluded. Representativeness is not claimed; the aim is to summarize recurring requirements across the CFP corpus analysis and to note where expectations vary between the sampled calls.

5.4 Recruitment, consent, and participation

Recruitment proceeded through brief invitations via existing networks, direct messages, for example WhatsApp, and a small number of referrals. To reduce intermediary-related pressure, invitations emphasized that participation was voluntary and that declining would not affect any relationship with the organization or individuals who facilitated contact. Before each interview, participants received clear information about the study and the conditions of participation, and consent was obtained and recorded in a format appropriate to the context, mainly verbal. This followed Brazilian ethical guidance for research in the human and social science. Participation was voluntary throughout. Participants could refuse to answer any question or withdraw at any time without consequences (Conselho Nacional de Saúde, 2016). Interviews were mainly conducted in person, with a small number held via WhatsApp phone call (project leaders) or zoom/google meet (specialists), when requested by the interviewee.

Procedural details recommended for transparent reporting were documented, including the method of approach, interview setting, whether any non-participants were present, and brief field notes written during and or after interviews (Tong et al., 2007). No recruitment was required for the CFPs. Documents were collected from official portals or sponsor websites and logged in the study’s document trail.

5.5 Data sources and procedures

Primary data consist of semi-structured interviews (see Appendix B) with the duration of approximately 30 minutes to 1 hour and 30 minutes, conducted in Portuguese. Topics included project aims, everyday tracking routines, experiences with supporter requests or reporting templates, local views of what counts as credible evidence, and feasible

adaptations to methods. Where participants consented, interviews were audio-recorded and transcribed, and key procedural details are reported transparently (Tong et al., 2007). Audio files were deleted after transcript verification and basic quality checks. Secondary data consist of routine project artefacts that reflect evidence practices in day-to-day delivery and were shared voluntarily by participants. These include indicator lists in use, attendance sheets, redacted WhatsApp screenshots, simple monitoring spreadsheets, and recent reporting templates (Baxter & Jack, 2008). Artefacts containing personal data were either received already redacted or redacted immediately upon receipt. Phone numbers, message IDs, and WhatsApp metadata were not copied into the dataset. Screenshots were redacted at source before storage, consistent with the study's ethics and data-protection approach (Conselho Nacional de Saúde, 2016; Câmara dos Deputados, 2018).

Documentary data comprise CFPs and associated rule materials, including selection criteria tables, mandatory document checklists, monitoring and reporting models, applicant guides, and annexes. A consolidated synopsis of the documentary corpus is provided in Appendix A, summarizing requirements across ten separate CFP documents/packages in aggregated form to reduce traceability and support anonymity of the field setting and participating initiatives. The appendix A synopsis was produced through AI-assisted summarization (using a large language model) and is intended as a transparency aid rather than a verbatim extract.

Specialist interviews followed a separate guide (see Appendix C) focused on the perceived realism of different impact-measurement tools, for example SROI and Theory of Change, in low-resource settings. The guide also covered the role of intermediaries and what a "minimal but fair" evaluation setup could look like. These insights are later used to interpret both project practices and CFP demands.

Because this study draws on multiple types of material (interviews, routine artefacts, and a documentary CFP corpus), it is useful to make explicit what each strand contributed and how it was used in the analysis. Table 1 summarizes the main data sources included in the study, their purpose in relation to the research questions, and where they are used in the thesis.

Data strand	What it includes (examples)	Purpose in this thesis	How it was used in analysis
Project leader interviews	13 semi-structured interviews (Portuguese), ~30–90 min; focused on impact meanings, routines, supporter/CFP experiences	Core evidence for **RQ1** (everyday evidence + local impact) and **RQ3** (fit/misfit)	Read in full; relevant passages extracted to structured notes; within-case summaries drafted; compared in cross-case matrix (Excel)
Specialist interviews	3 interviews with intermediaries/specialists on "realistic" tools and expectations in low-resource settings	Context for **RQ2**/**RQ3**; helps interpret how demands are understood and negotiated	Summarized into interpretation memos; used to clarify what "good enough" tends to mean in practice
Routine artefacts (shared by participants)	Attendance sheets, indicator lists in use, redacted WhatsApp screenshots, simple tracking spreadsheets, recent reporting templates	Triangulates interview accounts; shows what evidence looks like in delivery (**RQ1**)	Treated as illustrative traces; used to ground claims about routines and feasibility limits
Documentary corpus (CFPs)	10 CFP document packages + annexes/templates (criteria tables, mandatory documents, monitoring models, guides)	Defines the "official vocabulary" of evidence demands (**RQ2**) and provides comparison baseline for misfit (**RQ3**)	Extracted into a standard template (Excel); compared across documents by requirement type; synthesized into Appendix A

Table 1. Overview of study materials by data strand (own table).

5.7 Data management and protection

Audio files, transcripts, and shared artefacts are stored securely with restricted access. A de-identified key linking participant codes to identities is stored separately. Personal data are minimized, and direct identifiers are removed from quotations, consistent with ethical commitments to confidentiality and privacy in research with human participants (Conselho Nacional de Saúde, 2016). Data handling follows Brazil's data protection law, the Lei Geral de Proteção de Dados Pessoais (LGPD). In practice, this means collecting only what is necessary for the stated purpose, limiting access, and using safeguards to reduce the risk of unauthorized access, loss, or disclosure (Câmara dos Deputados, 2018). Working files are version-controlled, and retention periods are stated in the participant information sheet. Audio is deleted after verified transcription. Transcripts and de-identified artefacts are retained for up to 24 months post-examination, or as approved by

the ethics committee, and then securely destroyed. Only de-identified exemplar quotations are retained in the manuscript appendices.

Interviews are conducted and transcribed in Portuguese. Key excerpts used in the thesis are translated into English with AI assistance using de-identified text only, and are then manually checked by the author for meaning and tone. CFPs and related rule documents

are public materials and do not contain personal data of research participants. PDFs and extracted tables are stored separately, and source details and access dates are recorded in the audit trail.

5.8 Analytic strategy

Analysis was organized around the three research questions and the purpose–feasibility–legibility lens introduced earlier. Interview transcripts and field notes were read in full and revisited during the write-up. The interview material was read in full and revisited during the write-up. Early impressions and relevant passages were recorded in analytic notes, and interpretation moved back and forth between transcripts, notes, and drafting as the argument and comparisons across cases became clearer (Maguire & Delahunt, 2017).

The main analytic task was not to build a comprehensive codebook, but to identify and extract passages that spoke directly to the research questions, including how leaders describe “impact” in everyday terms, what forms of proof they rely on in delivery, and how they experience external demands in applications and reporting. To keep the analysis close to the material, structured analytic notes and case summaries were drafted using participants’ own wording and concrete examples. These summaries were then used to anchor the results chapter with contextualized quotations. Materials were organized in Excel files, including within-case summaries, cross-case matrices, and CFP extraction templates. To keep comparisons systematic while staying close to participants’ accounts, key extracts were summarized and organized in case-by-theme matrices, an approach aligned with the framework method’s emphasis on structured charting and comparison (Gale et al., 2013).

Within each case, short memos documented how evidence was produced and used in practice, which supporter or CFP demands were most salient, and what adaptations were described as feasible without undermining delivery. These memos were revised as additional interviews and artefacts were incorporated. Cross-case comparison was then

conducted through simple matrices that aligned cases under the same headings while retaining contextual detail. This provided a practical bridge between case-level accounts and the structure of the results, and it helped ensure that quotations were interpreted in context rather than treated as stand-alone findings.

The documentary strand was analyzed separately through structured extraction of each CFP package into a standard template, following the dimensions described in Section 5.5. This enabled transparent comparison across the corpus without constructing composite indices or assuming that requirements are comparable in their intent or enforceability. The focus was on what organizational profile and evidence practices the calls appear to presuppose, and which requirements are most consequential for small teams, particularly legal and administrative prerequisites, indicator and reporting expectations, and platform or reporting-cadence requirements.

Specialist interviews were treated as contextual interpretation rather than expert validation. They were used to clarify how common requirements are understood in practice, what types of evidence are typically considered "good enough," and which tools are seen as realistic or unrealistic in low-resource settings. Their role in the analysis was to support interpretation of recurring tensions by connecting project-side routines to the documentary expectations observed in the CFP corpus.

To support transparency, the analysis process was documented through a basic audit trail, including the CFP extraction templates, within-case analytic notes, and the cross-case comparison matrices. Descriptive counts (for example, how many cases mention a practice) are used sparingly to make patterns visible, without implying statistical inference. Sample adequacy is justified using information power, treating adequacy as dependent on the study aim, the specificity of participants, and the richness of the material, rather than on a fixed numeric target (Malterud et al., 2016). During fieldwork, recruitment decisions were informed by periodic checks of whether additional interviews were still adding substantively new insights relevant to the research questions (Hennink & Kaiser, 2022).

5.9 Researcher role and reflexivity

The researcher's position relative to the field is described because positionality shapes both how qualitative research is conducted and how accounts are understood and interpreted (Holmes, 2020). The report states who conducted each interview and

summarizes relevant interviewer characteristics, including credentials, occupation at the time of the study, gender, and prior experience or training in qualitative methods (Tong et al., 2007). It also notes whether any relationship existed prior to the study, and what participants knew about the researcher's reasons for conducting the research (Tong et al., 2007). Reflexive memos are kept throughout recruitment and analysis to surface assumptions, note surprises, and document decision points, and any prior relationships with specific projects are declared. All interviews are conducted by the researcher, which makes reflexivity particularly important.

5.10 Ethical considerations

Procedures follow Resolução CNS nº 510/2016 with respect to informed consent, the right to withdraw without penalty, and protections for privacy and confidentiality in research in the human and social sciences (Conselho Nacional de Saúde, 2016). Personal data handling follows the Lei Geral de Proteção de Dados Pessoais (LGPD), with processing grounded in a lawful basis and guided by principles including purpose limitation, necessity and minimization, security, and accountability, alongside de-identification wherever feasible (Câmara dos Deputados, 2018).

For WhatsApp audio interviews, consent is obtained before the interview begins and, where appropriate, recorded in audio form. Participants may opt out of recording and proceed with detailed note-taking (Conselho Nacional de Saúde, 2016). Participation may be discontinued at any time without prejudice.

Interview materials and research procedures were submitted to the FGV Comitê de Ética em Pesquisa (CEP) and received approval before any participant recruitment or data collection began. The documentary strand analyzes public CFPs and related annexes. No personal data of research participants are collected from these sources, and document quotations are kept brief, with sponsor and year cited.

5.11 Limitations of the design

This study uses a small, purposive (non-probabilistic) sample. This means it cannot support statistical generalization. Instead, the aim is to develop contextual, analytically grounded explanations that may be relevant to similar settings, rather than universal claims (Crowe et al., 2011).

A second limitation is that the depth of triangulation varies across cases. Routine artefacts were not equally available for all initiatives, so some case accounts rely more heavily on

interviews and field notes than others. This is mitigated by using multiple cases and drawing on multiple sources where available, and by keeping a transparent decision trail that documents what each interpretation is based on (Baxter & Jack, 2008; Nowell et al., 2017; Noble & Smith, 2015).

A third limitation concerns the interpretive nature of the analysis. The findings are based on the researcher's reading of interview material and documents. To make this interpretation open to scrutiny, the study documents key analytic decisions, keeps clear case summaries and comparison matrices, and uses direct participant quotes where appropriate so readers can see how claims were developed (Nowell et al., 2017; Noble & Smith, 2015).

Finally, two design choices require caution in interpretation. First, "small" is treated as a practical profile rather than a strict threshold, so cases close to more formalized NGOs are flagged and discussed carefully. Second, the documentary analysis reflects what CFPs formally ask for, meaning intended selection and compliance rules. These formal requirements may differ from how selection and reporting work in practice, so the documentary findings are interpreted descriptively, as indicators of what calls presuppose about applicants rather than as a full account of enacted processes.

5.12 Reporting structure and planned deliverables

The Results chapter (Chapter 6) is organized thematically rather than case by case. Section 6.1 briefly introduces the three strands of material (projects, specialists and CFPs) and the purpose–feasibility–legibility lens. Section 6.2 describes what "impact" looks like for the projects in their own terms, before turning in Section 6.3 to the evidence they generate in everyday practice. Section 6.4 summarizes what different types of CFPs ask for in terms of legal–administrative requirements, project justification, indicators, monitoring and reporting. Section 6.5 compares everyday evidence and CFP demands, highlighting areas of fit and persistent gaps. Section 6.6 sketches proportionate adaptations ("conversion recipes") that could help make existing evidence more legible without overburdening projects, drawing especially on specialist interviews. Section 6.7 then synthesizes findings back to the purpose–feasibility–legibility triangle and the three research questions.

Appendices include the project-leader and specialist interview guides, a brief description of each anonymized project, a list of the analyzed CFPs and annexes and the COREQ and

SRQR checklists. These materials are intended to strengthen transparency by showing how the study moved from design, through data, to the themes and interpretations presented in the Results and Discussion chapters (Tong et al., 2007; O'Brien et al., 2014).

6. Results

6.1 Reading the material

This chapter brings together three strands of material: interviews with leaders of small community-based projects in Rio de Janeiro, interviews with specialists in social impact evaluation and social projects, and a coded sample of Brazilian CFPs that are relevant for these initiatives. Read through the lens introduced in Chapter 2, the aim is to see how "purpose", "feasibility" and "legibility" come together in practice rather than as abstract categories.

The first strand consists of thirteen semi-structured interviews with leaders or coordinators of community projects, conducted mainly in Rocinha and Vidigal. The projects vary in activity and history: some are focused on sport and martial arts, others on dance, fashion and creative work, others on education, arts, environmental activities or support groups. Some have existed for more than a decade, others for a shorter period. Despite these differences, they share key features. They operate in territories marked by high social vulnerability and uneven public provision, they rely heavily on volunteer labor, and they manage their activities with very small budgets and minimal administrative support. Most have one or two central figures who combine coordination, teaching and fundraising in a single role.

The second strand consists of three specialist interviews. All three interviewees work in the field of evaluation and support of social projects in Brazil, in different institutional settings. One is an academic who writes about evaluation in the third sector and the construction of this field. The others work closer to practice, at the interface between universities, NGOs and community organizations. Together, they add an external vantage point on what counts as "good enough" evidence for small projects, which tools tend to be realistic or unrealistic in low-resource settings, and how broader funding and support arrangements shape which initiatives become visible.

The third strand is a sample of CFPs and related documents issued by public bodies, companies and philanthropic organizations. These are one of the main channels through which community projects access external resources, and they play an important role in defining what social projects should be able to present. The sample includes CFPs that explicitly mention community-based initiatives as well as more demanding public and corporate CFPs that leaders refer to in their interviews. This focuses on what these documents ask for in terms of objectives, activities, indicators and sources of verification, monitoring tools and platforms, administrative documents and certificates, narrative formats and reporting or feedback procedures.

The purpose–feasibility–legibility triangle introduced in Figure 1 provides a simple way of holding these pieces together. “Purpose” refers to what projects are trying to change in the lives of participants and in their territories. “Feasibility” refers to what they can realistically observe and document with the staff, time and tools they have. “Legibility” refers to how this evidence is translated – or not – into formats that external actors recognize as valid, from CFPs and reports to visits by funders or partners. The triangle is not applied as a scoring tool; it acts as a fit check when looking at how practices, demands and interpretations line up or diverge.

In practical terms, the chapter is organized thematically rather than by research question. Section 6.2 focuses on the projects’ own perspective, describing how leaders talk about the changes they see and which outcomes they prioritize. Section 6.3 stays with this perspective but shifts the emphasis to how these changes are observed, remembered and recorded in everyday practice. Section 6.4 then turns to the external side and summarizes what the coded CFPs ask for in terms of indicators, monitoring and documentation, and what kinds of organizational profiles these requirements presuppose. Section 6.5 brings these sides together, using the fit between purpose, feasibility and legibility to discuss where there is alignment, where there are systematic gaps, and what kinds of selection effects follow. Section 6.6 sketches proportionate adaptations that could help translate everyday evidence into more legible formats without overburdening projects. Finally, Section 6.7 synthesizes the findings through the lens of the triangle and links them back to the broader questions of the thesis. Throughout, insights from the specialist interviews are woven into the analysis, both to interpret the patterns that appear and to prepare the ground for the more normative discussion in the following chapter.

6.2 What “impact” looks like for the projects

Across the thirteen community-based projects, “impact” is not described primarily in terms of abstract indicators, but in terms of concrete changes in everyday life: how children and young people spend their time, how they relate to adults and peers, how families organize their routines, and how participants start to see their own future. When project leaders talk about results, they rarely mention numbers. They start with people: a child who “did not stay still” now managing to complete an activity, a mother who finally has somewhere structured to take her children every afternoon, an older woman who leaves the house to join a fitness class while her grandchildren train.

Projects differ in their main activities, some focus on sports and martial arts, others on arts and recycling, others on mixed educational and socio-emotional support, but their descriptions of impact converge around a few shared dimensions. This section organizes these dimensions into four clusters: (i) structured time and safe social spaces; (ii) discipline, focus and socio-emotional changes; (iii) family-level effects and intergenerational ties; and (iv) longer-term trajectories and opportunities. The aim is to show what “counting as change” means in the projects’ own terms, before turning to how they try to document these changes (Section 6.3) and how funders articulate their expectations (Section 6.4).

6.2.1 Structured time and a place to be

A first, very consistent layer of impact is for children to simply have a regular place to be after school. Project leaders repeatedly describe their work as creating structured time in the day and week: times when children and adolescents are not at home without activities, and not “out on the street”, but engaged in something that is predictable, organized and supervised.

Several initiatives run activities from Monday to Friday, often every afternoon; a few operate every day of the week, including weekends. One coordinator explains that her center has become a reference point “because we give support every day, from Sunday to Sunday”, with children coming daily and women sometimes staying for one or two months while they stabilize their lives and income. Another project leader describes how, in the afternoon, “they come every day, from Monday to Friday”, emphasizing that the project is not a one-off course but a continuous routine woven into daily life.

From the projects' perspective, this regularity is already a result. It means that children arrive at roughly the same times, learn the rhythm of sessions, and start to expect that the project will be open. Some coordinators use very simple signals to judge whether this is working: they notice children "waiting at the door" at the usual hour, or ask parents if there were days when the child did not want to come. When attendance remains high without being forced, this is interpreted as evidence that the project has become part of participants' routine, not just an occasional visit.

This structured time is also collective. Projects often share small spaces, hold classes in community courts or multi-purpose rooms, and receive children of different ages. One arts-and-recycling initiative, for example, grew from a family living room into a space where children of 5–14 years come every day to work with materials, paint and learn, supported by a small team of volunteer teachers. In a martial-arts project, different age groups rotate through the same physical area over the afternoon and evening, with younger children focusing on basic movements and older ones training more intensively.

In this sense, "impact" begins with the existence of a stable, recognizable collective space where children and families know they will be welcomed, recognized and engaged in something structured. Project leaders insist that this is not a minor achievement in contexts where many services are intermittent and where families juggle irregular work, long commutes and limited childcare support.

6.2.2 Discipline, focus and socio-emotional changes

A second group of outcomes relates to changes in behavior, attention and socio-emotional life. Almost all coordinators describe some version of the same narrative: children initially arrive "without firm discipline", "not able to concentrate" or "not used to rules", and over time they learn to stay in an activity, respect basic norms and relate differently to others.

In the martial-arts project, the teacher describes impact as "developing discipline, focus/concentration and motor coordination", not only for children without diagnosed difficulties but also for participants with disabilities. She mentions improvements that can be seen by eye: a child who could not coordinate basic movements gradually manages full sequences; a participant who barely spoke in class starts to interact and present themselves in front of the group. For her, a key sign of impact is when those changes are

also noticed by others, such as schoolteachers who report “more focus and discipline” in the classroom.

In other projects, leaders use different language but point to similar processes. The coordinator of the arts-and-recycling project talks about children who arrive “without anything to do” and end up assuming small responsibilities, taking care of materials and helping younger participants. At the multi-purpose center that supports both children and mothers, the founder emphasizes how children evolve from very young ages (from around 1.5 years) into older participants who stay for many years and eventually become trusted staff members. For her, seeing a former participant become an employee is a strong signal that the project has contributed not only to one-off behavior changes but to a longer process of socialization and learning.

Socio-emotional changes are more diffuse but equally central. Several coordinators speak about children becoming “more confident”, “more outgoing” or “less shy” as they participate in group activities, perform in presentations or manage to complete tasks they thought they could not do. Although these outcomes are hard to reduce to a single indicator, they appear repeatedly when leaders are asked what makes them feel that the project “is working”.

Taken together, these accounts describe impact less as a sudden transformation and more as a gradual accumulation of small behavioral and emotional shifts that become visible in day-to-day interactions: a child who now waits for their turn, completes a sequence, helps a peer, or addresses an adult with more confidence.

6.2.3 Families, caregivers and intergenerational effects

A third, very strong theme is that projects understand their impact as extending beyond the individual beneficiary to the family, especially to mothers and grandmothers. Many of the leaders explicitly say they think in terms of “the whole family, not just the child”.

This appears in different forms. In one capoeira-based initiative, for example, mothers do not simply sit and wait while children train. The project organizes a fitness class for them in parallel, so that while the children are in capoeira class, the mothers do a fitness session. For the coordinator, this is part of the project’s impact: mothers manage to leave the house, exercise, be with other people, and an older woman who became one of the first adult

participants is mentioned as someone who “loves this type of work” and has a relationship with the team that is different from that of a child or teenager.

In the multi-purpose center, the project explicitly combines activities for children with support to women, many of whom are mothers of children with disabilities or have experienced domestic violence. Women may sleep at the project for one or two months until they receive their first salary, have their own home and can continue with their life. The coordinator also prioritizes employing these mothers whenever possible, describing how some of them become part of the team. In this case, impact includes immediate safety and stability for women, as well as a pathway towards employment and a new routine.

In other projects, family effects are seen through parents’ and caregivers’ observations. Coordinators often mention that they “listen to parents and children”, asking what has changed at home. One leader explains that she pays attention when mothers say the child is not at home all the time anymore or when they ask why their son or daughter is not at the project on a given day – for her, this indicates that both child and family have incorporated the activity into their daily life and miss it when something interrupts the routine.

These examples suggest that, in the projects’ own view, impact is relational. It appears not only in the child’s behavior inside the activity, but also in how caregivers reorganize time, gain spaces for themselves, and perceive changes in everyday interactions at home.

6.2.4 Duration, continuity and trajectories

Finally, project leaders give considerable weight to time. Impact is not described as something that can be fully captured in a short intervention cycle. Instead, they stress longevity and continuity: how long a child stays, whether families return, and whether participants move into new roles.

Several coordinators remark that children remain in the project for many years. One mentions participants who arrived from around one year of age and are now ten, reflecting her longer experience in education more broadly and the fact that the project itself has been running since 2017. Another emphasizes that some girls stayed until they were fifteen or sixteen and later became employees, closing a full circle from beneficiary to team member.

Time is also reflected in the way projects see their own calendar. Initiatives that began informally, for example inside a family home, evolved into year-round activities rather than limited “project periods”. One coordinator recalls how her father broke a wall in their house to make space for children to work with recycled materials so they would not be standing still in the street without any activity, and that from the beginning it was the whole year. For her, this uninterrupted presence in the community is part of the impact: it signals commitment and makes it possible to accompany children through different life stages.

In some cases, duration is also about intensity. The multi-purpose center, for instance, supports children daily and offers temporary residence to women until they reach a certain level of autonomy. The capoeira project describes a routine in which children grow from very small participants who mainly play and familiarize themselves with the space into older adolescents who train more seriously, while older adults and mothers join their own activities in parallel.

Taken together, this section shows how project leaders first define impact in substantive terms: structured time, behavioral and socio-emotional shifts, family-level changes and long trajectories in the same initiative. The next section stays with the projects’ perspective but shifts the focus from what counts as change to how these changes are observed, remembered and recorded in everyday practice.

6.3 Evidence-in-practice: how projects make change visible

This section shifts from *what* changes to *how* these changes are made visible in everyday practice. Across the studied cases, evidence-in-practice is built less through stand-alone evaluations and more through a bundle of everyday signals, lists, photos, archives and occasional reports that coordinators can maintain alongside delivery work; Section 6.4 then contrasts these routines with the documentary demands set out in the CFP corpus (summarized in Appendix A).

Four layers appear repeatedly: (i) everyday signals and conversations; (ii) simple lists, counts and micro-monitoring routines; (iii) photos, videos and other digital traces; and (iv) archives, cadence and roles in documentation. Together they show that projects are not “empty” of evidence, even when they rarely use formal frameworks or specialized tools. Specialist interviews describe this bundle of lists, stories and partner feedback as a

form of "everyday evidence" that is often undervalued when assessment is discussed only in terms of formal frameworks.

6.3.1 Everyday signals and conversations

A first layer of evidence is made of small, recurrent signals that coordinators read as signs that the work is having an effect. Many of these mirror the outcome themes in Section 6.2, but here the emphasis is on how they are *observed* and interpreted rather than on their intrinsic value.

Punctuality and routine are central examples. Leaders describe children who arrive well before sessions start and parents who complain when an activity is cancelled because "the child is already waiting at the door". In a football project that has existed for decades, this insistence on coming, even in bad weather or when communication fails, is treated as proof that the activity has become part of children's weekly rhythm. Similarly, in a capoeira project with multiple training spots, parents call when one session does not happen, asking why there was no class. These reactions are not written in a spreadsheet, but they function as everyday indicators: coordinators notice when children do not want to miss the activity and when families reorganize routines around it.

Behavioral and socio-emotional shifts are tracked in a similar way. Leaders pay attention to how children behave in the classroom or training space: whether they can now wait for their turn, stay focused for longer, or help peers. They also listen for external confirmation. In some projects, school teachers report "more focus and discipline" in the classroom as well as fewer absences on days of the project, in others, caregivers comment that children are calmer, more respectful, help more at home and are more motivated to do homework. Those comments, usually shared in passing during drop-off or pick-up, are stored in memory and sometimes in short notes. They anchor leaders' sense that change is happening, even if they are not translated into formal scales.

For projects that work with circular economy and environmental education, everyday signals take other forms. Coordinators notice when families start to treat reuse and second-hand consumption as "normal", when artisans begin to earn small incomes from products made with recycled materials, and when fairs consistently bring together dozens of stalls and families. These observations are often summarized in simple phrases such as "thirty families generating income" or "a fair full of stands with reused materials". Even

when counts are approximate, they provide a tangible frame for stories about changing consumption and local economies.

Taken together, these examples show that a substantial part of evidence-in-practice is relational and conversational. Coordinators watch how people behave, listen to how they talk about the project, and store these observations as a mental record of change. Only some of this makes its way into written documents.

6.3.2 Lists, simple counts and micro-monitoring routines

Alongside everyday signals, almost all projects maintain some form of list or simple count. These tools are modest in format but central for organizing activities and, when needed, for showing that work took place.

Attendance and registration lists are the most common. They typically include names, contact details, age, and presence by day or week. Some projects also mark whether the child is enrolled in school or has special needs. In the football initiative, for instance, all participants must be matriculated in public school; the list doubles as a way to check this and to monitor how many children are attending regularly. In the reinforcement and shelter project, separate lists are kept for children, mothers and, where relevant, information related to disabilities, so that support can be tailored.

Beyond basic attendance, several projects have developed small monitoring loops that are integrated into daily practice. A Ju-jitsu project, for example, conducts an initial conversation with each child and caregiver when they first come to the project, asking about school, routines at home and any specific difficulties. During classes, the coordinator informally follows up: whether the child is still enrolled, how they are doing in class, if there have been conflicts at home. These notes may stay in a notebook or simply in the coordinator’s head, but they function as a structured way to detect change in school engagement, family dynamics and behavior over time.

In a cultural project, the coordinator pays attention to punctuality, discipline during tasks and feedback from parents about behavior at home. When more serious emotional situations arise, the team may involve a psychologist linked to the project. Again, there is no formal scale, but simple registers and regular conversations create a pattern of information that is revisited when evaluating how children and families are progressing.

Some initiatives also keep approximate counts related to their specific focus. The circular-economy project, for instance, tracks how many families participate in a fair, how many stands are set up, and, in some cases, how many items of school material or clothing were exchanged or donated. These numbers are not always collected with the precision of a survey, but they provide anchors for speaking about reach and intensity.

From a feasibility perspective, these tools are light: they use notebooks, simple tables or basic spreadsheets. From a legibility perspective, they already resemble the minimal indicators many funders ask for (number of participants, frequency, type of activity), even if projects do not always present them in that language.

6.3.3 Photos, videos and digital traces

A third layer of evidence comes from visual and digital records. Coordinators often say that "everything is on the phone". Photos and videos document everyday activities, special events and participants' trajectories, and are used both for memory and for accountability when partners request proof of delivery.

Several leaders recount situations where funders asked for photos, frequency reports and receipts as conditions for continued support. In response, projects sent combinations of attendance lists, brief narratives of activities and impact, and images of sessions, fairs or presentations. In other cases, project leaders proactively share photos and short updates with donors, local partners or on social media, anticipating that this material may help when applying for future funding.

Visual traces also play an internal role. In creative and sports projects, murals painted by participants, trophies from tournaments, costumes from shows or posters from past events remain on the walls or in storage. They act as reminders of what has happened, who took part and how the project has evolved. Leaders use them when explaining the initiative to visitors, parents or potential supporters, pointing to specific murals, medals or photographs as concrete examples of long-term impact.

Digital platforms extend this archive. The fashion and empowerment project, for example, has accumulated boxes of newspaper and magazine clippings, printed photos and invitations since the 1990s, and now adds videos and coverage on online platforms. The coordinator explains that this is both a way of not forgetting ideas and a way of having something to show when a call or partnership appears. Over time, this becomes a

longitudinal record of trajectories: which young people walked in fashion shows, which events took place in which cities, and which institutions recognized the project.

These visual and digital traces are rarely labelled as “monitoring”, but in practice they provide a rich, time-stamped record that can be mobilized as evidence when needed.

6.3.4 Archives, cadence and roles

Looking across cases, it is possible to see not only what is documented, but also how often, by whom and where.

Monitoring routines tend to happen around key activities and funder reporting dates, rather than on a fixed schedule. Attendance lists are usually updated daily or weekly, in the classroom, court or training room. Stories and reflections are written down less frequently, often when a report is requested or when coordinators decide to summarize a period for their own use. In projects that have links to companies or municipal programs, there is often a quarterly or annual reporting cycle in which leaders compile lists, receipts and brief narratives of results. When there is no external request, documentation still happens, but in a more ad hoc way: photos of a particularly meaningful moment, a Facebook or Instagram post about an event, or notes in a personal notebook.

Roles in evidence production are pragmatic and shaped by team size. In many projects, the main coordinator does almost everything: designs activities, teaches or supervises, keeps attendance, talks to families, and writes reports when necessary. Where there is a slightly larger team, tasks are distributed informally according to skills and preferences. Someone who is “good with numbers” helps with spreadsheets; someone who enjoys writing drafts posts or short reports; volunteers or younger participants sometimes support with photos and social media. Some organizations have occasional support from psychologists, pedagogues or volunteers with administrative experience, but this is not the norm.

Storage is fragmented but functional. Part of the documentation stays in physical notebooks and folders at the project site; another part sits on personal phones and laptops, especially photos and videos; and when a partner requires it, some information is uploaded onto external platforms. Coordinators are aware that this dispersion creates risks: a lost phone, a broken computer or a volunteer who leaves can mean that some

material disappears. At the same time, they have limited time and infrastructure to centralize everything in a single system.

Taken together, these patterns suggest different “depths” of documentation within a small sample. Some projects already operate what could be called a light monitoring system, combining intake conversations, periodic check-ins, simple lists and photo trails. Others have decades of material in paper archives and online but activate it mainly when a call or public presentation demands a more systematic narrative. A third group relies heavily on lived memory and external recognition, documenting less but still accumulating strong stories and reputational signals.

From a purpose–feasibility–legibility perspective, the picture is of initiatives that already produce significant evidence, but in formats shaped by their capacity and relationships. Feasibility clearly constrains what can be formalized: tools must fit around teaching, care and informal support. At the same time, many of these everyday artefacts – lists, counts, stories, photos and archives – are not far from what external actors recognize as indicators and documentation.

6.4 What CFPs and funders ask for

Up to this point, the focus has been on how projects themselves understand change and how they document it in everyday practice. This section turns to the other side of the relationship: what different types of funders, especially public CFPs and corporate programs, ask organizations to present in order to be selected and to keep support. The material combines the sample of CFPs and guides analyzed with the project leaders’ accounts of past funding cycles and attempted applications. To protect anonymity, the CFPs and institutions are referred to in generic terms (for example, “a state socioenvironmental call” or “a corporate sports program”), and a consolidated synopsis of requirements across the ten CFP document packages analyzed is provided in Appendix A. Taken together, the documents and interview accounts show that demands operate on at least four layers: basic legal–administrative eligibility; project description and justification; planning, indicators and monitoring; and ongoing reporting and accountability. In some CFPs there is also an explicit layer of expectations regarding the composition and qualifications of the team, such as minimum years of experience and higher education degrees. Selection criteria cut across these layers and reward organizations that are older, more formalized and able to write in the language of the CFP.

6.4.1 Legal and administrative eligibility

The first layer consists of basic eligibility conditions that must be met before any substantive analysis of the project takes place. Several public and corporate CFPs require that the proposing organization is a legal entity with a CNPJ, usually with a minimum number of years of existence at the time of application. One large socioenvironmental CFP, for instance, demands proof that the organization has existed for at least three years, through a recent CNPJ certificate, and that it can submit its founding statute or equivalent registration document. A state-level program adds that the organization must present balance sheets and income statements for the last two financial years, signed by a certified accountant, as well as multiple "negative debt certificates" for federal taxes, labor liabilities and FGTS. Another annex lists compulsory registration in relevant public councils (for the rights of children and adolescents, social assistance or education) as part of the technical qualification package. In the same set of documents, organizations are encouraged to present annual activity reports and, where available, external audit reports as signs of "transparency" and adequate management control. This presupposes a level of formalization and professional support that goes beyond basic bookkeeping, a pattern that recurs across the CFP corpus (see Appendix A).

A corporate CFPs follow a similar pattern, even when the projects themselves are cultural or sports initiatives. A corporate culture–sports call, for example, asks organizations to upload a CNPJ card, statutes or contract, proof of an operating license, and declarations signed by the legal representative. It also requires evidence that the project is already approved, or at least under analysis, in the relevant public incentive system, with a protocol number generated by the federal platform. In practice, this means that many projects would first need to navigate a separate approval process in the incentive law system before they can even be considered eligible for the corporate call.

In parallel, an international microgrant form aimed at small projects does not require Brazilian certificates, but still expects a formal organizational identity. Applicants are asked to describe the organization, previous activities and how they are financed, and to indicate whether there is a partner organization involved. Even where the legal bar is somewhat lower, the underlying assumption remains that there is a registered entity with someone responsible for accounting and for signing the agreement.

For community initiatives that started informally and only later obtained a CNPJ, these requirements are often described by coordinators as the first exclusion point. Several interviewees explain that they had to decline or could not complete applications because the space used by the project did not have regularized documentation, or because they did not yet have an accountant, financial statements or council registrations in place. In their accounts, “being legal” is not just about having a CNPJ; it often means having a fully regularized institutional and physical infrastructure that goes beyond their reach. When CFPs also require a history of audited accounts, multiple negative certificates and formal council registrations, the distance between these expectations and the realities of small community projects becomes even clearer.

6.4.2 Project description, problem framing and justification

Once basic eligibility is met, CFPs move to detailed project description. Here, they tend to ask for long narrative fields in which organizations must describe their trajectory, the project and its context in a formal, technical language.

Corporate and state socioenvironmental programs typically require applicants to explain why the organization was created, what challenges it has faced and what main achievements it can point to. They ask for descriptions of other projects currently run by the organization and the results achieved. For the specific project being submitted, forms ask for a concise summary, an explanation of the main objectives and a clear definition of the target group and territory, often with explicit emphasis on children, adolescents and young people in situations of risk or vulnerability.

One of the larger socioenvironmental CFP goes further and incorporates these elements into its scoring grid. Projects receive points for how clearly they describe the problem they seek to address, how well they justify the proposed intervention and whether they present socio-economic data on the area of operation. The maximum score is attributed to proposals that present an “excellent” justification and consistent socio-economic information about the municipality, region or neighborhood; intermediate scores are given to those with good but less detailed justification; and proposals with only minimal justification and sparse data receive the minimum score. Those that fail to describe the problem and its context are eliminated.

The international microgrant form follows the same logic in a simpler way. It asks for a short description of the situation that motivates the project, the main objectives and why

the target group has been chosen. It also requires applicants to state whether there are risks or unintended side effects and how these will be mitigated.

Across all these documents, there is a clear expectation that organizations will frame their work in terms of a well-defined problem, justify their approach using contextual information, and show that they understand the broader social or environmental landscape in which they operate. The ability to produce this kind of written justification is treated as an indicator of organizational maturity. For projects whose coordinators are more accustomed to explaining their work orally, in community meetings and everyday conversations, this expectation of long, technically written justifications represent an additional threshold that is not directly related to the quality of local practice.

6.4.3 Workplans, methods and indicators

A third layer concerns the internal logic of the project: how objectives, activities, methods and expected results fit together, and how progress will be monitored. Most CFPs require a detailed workplan, sometimes broken down into activities, responsible people, timelines and locations. They also ask applicants to explain their methodology and to link it explicitly to expected results.

In the state socioenvironmental call, for example, one section of the selection criteria focuses on the connection between objectives, methodology, activities and expected results. Projects that show “excellent connection” between these elements receive the highest scores, with lower scores assigned to those where the links are only partially clear. Proposals that do not make these connections visible are eliminated. Another part of the same table stipulates that projects must present clear objectives and a detailed description of activities; again, failure to do so leads to automatic exclusion.

Corporate CFPs adopt a similar approach. The culture–sports CFPs analyzed asks applicants to describe the main activities and the results expected for each of them. It then explicitly requires them to define indicators that will be used to measure the success of the project and to explain how each indicator will be verified and with what periodicity. The form suggests that indicators should be measurable, gives examples of possible verification sources, and asks whether indicators will be checked monthly, bimonthly or at other intervals. In a large sports program, one of the selection criteria specifically values organizations that “demonstrate knowledge of data collection and use of indicators”,

signaling that familiarity with monitoring tools is itself treated as a sign of technical capacity.

The international microgrant form includes a specific field on indicators and effects. Applicants are asked to state which quantitative and qualitative criteria will be used to demonstrate the success of the project, what changes they expect to see and how the durability of results will be ensured. The same form asks whether an external evaluation is planned, indicating that, even in smaller schemes, funders often imagine evaluation in terms of formal designs rather than only internal learning processes.

In the project leaders' accounts, this layer appears when they talk about what was requested by large corporate supporters or by public cultural CFPs. A coordinator who received support from a major company explains that, in addition to basic information about the project, she had to produce a description of activities, specify the number of participants, and, after the project, report on whether the planned events had taken place. In another case, a leader recalls that a municipal CFP asked for a project description with objectives, justification and a simple result framework, even though the reporting later focused more on financial accountability and photographic documentation.

Overall, across different types of CFPs, there is a shared expectation that organizations will operate with something close to a "results chain": clear objectives, a set of planned activities, a defined target group, and a list of indicators and verification practices that show whether the project is on track. In one socioenvironmental call, this appears explicitly in a selection item that asks whether the project "provides instruments of monitoring/accompaniment of the activities to be developed", with higher scores awarded to proposals that already indicate how they will track implementation. In practice, this expectation moves much closer to the logic of formal evaluation frameworks discussed by specialists, even when the organizations targeted by the CFPs work with very small teams and rely on everyday evidence such as attendance lists, stories and partner feedback.

6.4.4 Budgets, financial control and accountability

Budgeting and financial control form a fourth layer of demands. All CFPs analyzed require a detailed budget, often separated into categories such as personnel, services, equipment, materials, communication and administrative costs. Applicants must assign values to each item and justify why the expense is necessary for the project.

In the state socioenvironmental call, the selection criteria include a specific item on the compatibility between the proposed activities and the budget. Projects that present a consistent and well-justified budget receive maximum points. The international microgrant form also demands that all expenses be broken down and specified. Corporate programs go further by asking proponents to classify costs according to the rules of public incentive laws, separate personnel from operating costs and, in some cases, indicate the form of contracting for each position (CLT, service provider, volunteer).

These expectations continue after selection. Corporate and public funders alike require detailed financial reporting, with invoices and receipts for all expenditures. One project leader who received support from a large company describes how she has to compile a report containing the names of participants, records of who entered and left the project, and all fiscal notes related to the funds used, alongside a bank statement showing the flow of resources. In her words, “you have to have a financial report of income and exit, justifying what it is being used for”. Another coordinator emphasizes that, even when final narrative reports are short, supporters expect clear documentation of how money was spent and sometimes conduct site visits to verify that activities are taking place as described.

The underlying assumption is that organizations have internal systems capable of tracking expenditures per budget line and producing itemized financial reports, which in turn presupposes an accountant or at least someone responsible for financial control. For many of the small projects interviewed, who rely on simple cashbooks and volunteer support, this level of financial structuring is only partially present, which limits their ability to apply for and manage larger, more regulated funding lines.

6.4.5 Monitoring, reporting and platforms

A fifth layer concerns how projects will be monitored and how results will be reported over time. Some programs make this explicit by presenting monitoring templates or asking for a specific monitoring and evaluation plan. One corporate sports program, for instance, provides a separate monitoring spreadsheet as part of its documentation, which suggests that selected organizations will be required to report periodically against predefined indicators and targets; Appendix A summarizes how reporting cadence, verification expectations, and template use recur across the documentary corpus.

Several CFPS are operated through online platforms where organizations must create a profile, upload documents and fill in electronic forms. These systems typically generate an automatic confirmation once the application is submitted and are later used to upload reports and supporting documents. The corporate culture–sports call analyzed provides a step-by-step guide on how to register the organization, fill in the project information and upload the required files, highlighting that incomplete documentation at the time of submission can lead to elimination in the second phase of analysis. In another case, the main coordinator is required to record and upload a short video presenting the project, which introduces demands not only in terms of internet access but also of digital self-presentation.

Public socioenvironmental programs also frame monitoring in terms of continuous data collection. One call mentions that monitoring will be conducted through the analysis of indicators associated with project objectives and that all selected projects will be subject to ongoing follow-up and evaluation by internal teams. The same document refers to intermediate and final reports, without spelling out their exact format in the excerpt examined, but signaling that monitoring is treated as part of the contractual obligations.

In contrast, the international microgrant form is less prescriptive but still expects a description of how the project will ensure the durability of results and hints that a brief report will be requested at the end of the funding period.

From the project leaders' perspective, these requirements translate into periodic obligations to send attendance lists, photos, brief narratives and financial statements to funders. Some emphasize that corporate supporters schedule visits and contact them in advance to see activities on site. Others recall that a cultural call required only a brief final report and photographic records, but not a complex monitoring plan. Even in lighter cases, however, the expectation is that organizations can produce documentation on demand. For projects with limited internet access, shared computers and coordinators who already combine multiple roles, learning to use platforms, uploading documents and keeping track of deadlines become additional tasks that compete with direct work with participants.

6.4.6 Selection criteria and competition

Beyond the layers described above, some CFPs make their selection criteria explicit in scoring tables that show how projects will be compared. A large socioenvironmental call,

for instance, assigns points to organizations based on their years of existence, number of technical capacity attestations and history of partnerships. The maximum scores are reserved for organizations with more than ten or twenty years of operation, five or more attestations of capacity, and several documented partnerships with public or private entities. Organizations with only one attestation or fewer years of experience receive significantly lower scores, and those with no attestations at all receive zero on that item. The same table includes a separate criterion for partnerships, stating that it is “desirable” for the organization to evidence partnerships formed with other institutions over the last ten years. Extra points are awarded to projects that can attach extracts of contracts, cooperation agreements, partnership terms or declarations from partners, which presupposes a history of formalized relationships and written records. The same table attributes more points to projects that reach a larger number of direct beneficiaries, which favors mass activities over small, intensive formats such as shelters or disability support.

Other items in the same table score the quality of the workplan, the strength of the justification, the consistency of the connection between objectives, methodology, activities and expected results, the size of the target group and the degree to which the project incorporates equity dimensions such as gender and race. One criterion asks whether the project presents “innovative elements” or “differentiated social technologies”, awarding additional points to proposals that can describe their approach in these terms. Innovation is treated as an additional positive criterion, considered a “differential” for projects that present new approaches within the call’s thematic scope. For leaders who did not go through higher education and who are used to speaking about their work in everyday language, the notion of “innovation” as a technical category can be difficult to translate. This makes it easier for organizations with staff trained in policy or project jargon, often led from outside the community, to frame their practices as innovation and to capture these extra points, even when community-based initiatives are also experimenting with new combinations of activities and support.

In at least one sports-related call, there are also explicit points for the composition and educational level of the team. Projects receive higher scores when a greater number of board members or staff have a university degree in sports or a related area, while initiatives without any formal higher education in the team receive zero on this item. Other criteria reward the number of programs previously executed by the organization. Together, these rules tend to favor organizations that already operate as professional

NGOs with specialized staff, formal networks and a portfolio of past projects. Corporate CFPs, while not always providing as detailed a scoring table in their public guides, describe a multi-stage analysis process. Typically, there is an initial eligibility check, a documental analysis and then a merit analysis where projects are assessed on coherence, alignment with thematic priorities, potential for impact, feasibility and the organization's track record.

For small community-based projects, these criteria mean that they are competing not only on the substance of their work but also on organizational age, documentation of previous experience, number of formal partnerships and the polish of written submissions. Several leaders describe having "reached the final stage" or "almost passed" in competitions but not understanding where their proposals fell short, since they do not receive feedback from the selection committees.

Taken together, the CFPs analyzed present a picture of a funding environment in which being eligible and competitive requires more than everyday evidence of change. It requires a set of legal, administrative and narrative capacities that many grassroots initiatives only partially possess. Criteria that reward long histories of formal partnerships, the ability to describe "innovative" social technologies and the existence of pre-defined monitoring instruments systematically favor organizations that already operate with professionalized teams, dense institutional networks and experience in writing for CFPS. For projects led by residents of the communities themselves, whose main assets are local legitimacy, continuity and everyday evidence, these expectations are much harder to meet, even when their substantive work is tightly aligned with the stated priorities of the CFPs.

6.5 Fit–gap between everyday evidence and CFP demands

The previous sections have shown, in parallel, how small community projects build and use evidence in daily practice and how funders and CFPs define what counts as sufficient proof of capacity and impact. This section brings these two pictures together. Rather than treating them as separate worlds, it examines where there is alignment between the evidence that projects already generate and what a "typical" call asks for, and where there are gaps that make it difficult for these initiatives to be selected or to sustain funding.

For clarity, the comparison is organized around five dimensions: basic descriptions of activities and beneficiaries; quantitative counts and indicators; monitoring routines and archives; legal–administrative and financial requirements; and written form and language.

6.5.1 Activities, target groups and substantive focus

On the most basic level, there is broad alignment between projects' substantive focus and the thematic priorities of the CFPs analyzed. Many of the them explicitly seek projects in areas such as education, sports, culture, socioenvironmental protection and rights promotion. They prioritize children, adolescents and young people in situations of vulnerability and, in some cases, highlight the importance of working with girls and women, people of color and indigenous populations, people with disabilities and other underrepresented groups.

The projects in this study operate precisely in these domains. They run football and martial arts classes, arts and recycling workshops, dance and cultural activities, educational support, and combined shelter and tutoring for women and children. Their target groups are children and adolescents from communities with uneven public services, precarious infrastructure and everyday exposure to risk. When coordinators describe their work, they situate it in relation to these conditions and focus on creating structured time, safer spaces and opportunities for learning and coexistence.

In this sense, there is a clear "fit" in terms of what problems are being addressed and who is being served. Projects and CFPs share a concern with inclusion, education, sport and culture for socially vulnerable groups. In theory, this should make them natural candidates for one another.

6.5.2 Counts, indicators and signs of change

A second dimension concerns how change is registered. As shown earlier, everyday practice already involves a range of quantitative and qualitative tools: participant lists, attendance sheets, simple counts of children, families and events, photos, videos, stories of individual trajectories and informal feedback from parents and teachers. Some projects also keep basic registers of age, school and contact details, and a few have begun to track trajectories more systematically over several years.

CFPs typically ask for the number of beneficiaries, often with minimum thresholds, and for some indication of expected results and indicators. Corporate and public programs demand that applicants estimate how many people will be reached, how often and in what

way. Some require that indicators be measurable and that verification sources and frequencies be specified; others request that applicants state how they will know whether the project has succeeded.

There is partial alignment here. The counts that projects already make – how many children are attending, how often they come, how many families participate in specific activities – respond directly to the numeric demands of CFPs. When leaders speak of children arriving on time, staying throughout the year, improving at school or becoming monitors, they are in effect pointing to outcomes that could be turned into simple indicators.

The gap lies in translation and consistency. Everyday evidence is organized primarily for internal and relational purposes: to know who is present, to plan activities, to share good news with families and partners. It is not always stored in a form that can be easily transferred into CFP templates. For example, attendance lists may be kept on paper in binders, photos may be stored on personal phones without dates, and stories may be held in memory or in dispersed messages. In addition, baselines are rarely formalized: projects know how children were "before" mainly through conversations and observations, not through systematic initial registrations.

From an CFPs perspective, this makes evidence appear fragmented. The information needed to define indicators and to show change over time is present but not assembled into the specific formats requested: a table with targets and verification methods, a brief monitoring plan, or a set of before-and-after numbers that can be summarized in a report.

6.5.3 Monitoring routines and archives

A third dimension is the frequency and depth of monitoring. In daily practice, projects already operate with routines: attendance is updated each session or week, check-ins with families and teachers happen periodically, and when there is a partner or supporter, some form of reporting is prepared, even if short. Over the years, some initiatives accumulate archives of photos, press clippings, trophies and murals that testify to sustained activity and to participants' trajectories.

CFPs, especially the larger ones and corporate programs, assume and demand a more formalized monitoring structure. They expect organizations to commit to periodic reporting cycles, to track indicators according to a defined schedule, and to feed

information into online platforms or standardized spreadsheets. Monitoring is presented not only as a way to account for resources but also as part of how impact is assessed across a portfolio of projects.

Here, too, there is a partial match. The rhythm of activities and basic monitoring in projects corresponds in broad terms to what funders hope to see. The gap lies in the level of standardization and in the ability to sustain documentation over time under contractual obligations. A small football or arts project that updates a notebook after each class and occasionally compiles numbers for a supporter is operating at a different scale of administrative effort from what is implicitly expected by a multi-year call that assumes dedicated monitoring staff.

Specialists interviewed in this study point out that this misalignment is not due to a lack of disposition on the part of projects, but to the absence of time, training and funding for evaluation work. From their perspective, everyday monitoring routines are a reasonable starting point, and demands become problematic when they presume research-like infrastructures that small teams do not have.

6.5.4 Legal–administrative and financial prerequisites

The sharpest gaps appear in the legal–administrative and financial dimensions. While projects have taken steps towards formalization – several now have CNPJ, some have a basic statute, and many maintain a bank account and simple financial records – the level of documentation required by many CFPs goes much further.

As detailed in Section 6.4 and synthesized across the documentary corpus in Appendix A, large public and corporate calls require several forms of organizational proof. These commonly include a history of at least two or three years of formal existence, complete financial statements for recent years signed by an accountant, multiple negative debt certificates, registration in thematic councils, and technical-capacity attestations for both the organization and the project coordinator. In addition, some expect that the project itself is already approved in a public incentive law system and that property and operating licenses are regularized.

Project leaders frequently describe these requirements as the point at which their attempts to apply break down. Even those who have managed to register a CNPJ within the last few years struggle to produce audited financial statements or to regularize the physical

spaces they use, since community buildings often lack formal documentation. Others emphasize that they would need a lawyer and an accountant to navigate the process and that they do not have the resources to hire such services. For small initiatives that have grown organically from voluntary efforts, this layer of demands is experienced less as a question of impact and more as a barrier to entry.

Financial reporting expectations amplify this imbalance. While projects do keep receipts and note expenses when they receive support, the capacity to produce detailed, line-by-line financial reports in the formats required by corporate or public funders is uneven. Those that have had access to intermediaries or partnerships sometimes manage to meet these standards; others rely on simpler records that are sufficient for local supporters but would not pass more rigorous audits.

6.5.5 Written form, language and feedback

Finally, there is a dimension of form and language. CFPs assume that organizations can express their work in a certain style: they must write in standard Portuguese, use the vocabulary of objectives, indicators, methodologies and socio-economic data, and fill long narrative fields within character limits. Selection criteria reward proposals that present “excellent detailing” or “good justification” and penalize those that are vague or poorly structured.

Project leaders often highlight this as a specific difficulty. Some say explicitly that they would need “someone who understands this CFPs language” to help them write, and that they do not have the time to learn and to follow each step. Others recount experiences of sending stories, photos and links to social media, along with simple reports and justifications, and being told that they “almost passed” without knowing what was missing. The absence of feedback from unsuccessful applications reinforces the sense that the rules of legibility are unclear.

From a fit–gap perspective, the content of what projects have to say matches many of the concerns of funders: they can describe what they do, who they work with, what changes they see and why their initiative is needed. The misfit lies in the expectation that these narratives be pre-formatted to match CFPs templates and in the lack of support to translate between everyday language and technical terms.

Overall, this comparison suggests that the main misalignments do not lie in the absence of impact or of evidence in community projects, but in the way institutional demands are structured. The triangle of purpose, feasibility and legibility helps to make this visible. Everyday evidence is purpose-driven and feasible; call demands are legibility-driven and often exceed feasibility for small teams. The next section explores what kinds of proportionate adaptations could bridge this gap without overburdening projects.

6.5.6 Specialist views on formal measurement frameworks

Specialist interviews add another layer to this fit–gap picture by situating everyday practices and CFP demands in relation to formal impact measurement frameworks. All three interviewees are familiar with tools such as social return on investment (SROI), cost–benefit analysis, logical frameworks and theories of change. Their accounts converge on two points: first, that some of these tools are not realistic for small, community-based projects; second, that it remains important for such projects to measure something, rather than rely only on general impressions.

On the more demanding end of the spectrum, specialists describe SROI and cost–benefit approaches as resource-intensive and technically demanding. To produce credible monetary estimates of impact, they argue, organizations would need baseline data, clear comparison groups or counterfactuals, sufficient sample sizes, stable implementation over time and professional support in costing outcomes. In their view, these conditions rarely hold for community-based initiatives with one or two coordinators and no dedicated evaluation staff. When SROI-style calculations are attempted without these foundations, the resulting numbers are seen as fragile and potentially misleading. Rather than strengthening credibility, they can raise questions about over-precise figures that are not well supported by data.

The same concern appears in their discussion of strong causal claims more generally. One specialist notes that it is tempting to attribute improvements such as better school grades or increased discipline solely to participation in a project, especially when leaders see large changes in specific children. However, from an evaluation perspective, they emphasize that such outcomes are influenced by many factors: family support, school quality, broader social policies and individual trajectories. To claim that a given project “caused” these changes in a strict sense would require control groups or other robust designs and sustained data collection, which are usually beyond the reach of small

initiatives. In this sense, pushing community projects towards heavy frameworks or strong attribution claims without the necessary conditions can be “hurting” rather than helping, because it encourages them to make promises that are hard to defend.

At the same time, specialists are cautious about simply abandoning measurement. They argue that communities, participants and funders have a legitimate interest in knowing whether a project is achieving what it set out to do. From their perspective, the question is not whether to measure, but how to do so in ways that are proportionate to capacity and appropriate to the kind of change involved. Here they converge on a more modest family of tools: simple theory of change and short results chains.

All three specialists point to theory of change as a realistic and useful option for small projects, provided they are kept simple and developed with support. In their view, a theory of change that fits community initiatives would not be a large, donor-driven diagram, but a concise narrative or visual that makes explicit the links between activities, immediate outputs and a small set of short- and medium-term outcomes. For example, football training three times a week, combined with conversations about school, might plausibly contribute to more structured routines, stronger bonds with adults and, over time, better school engagement. The theory of change would describe these links in everyday language, and then be used as a reference point for choosing a few feasible indicators and for selecting stories that illustrate different steps in the chain.

Specialists stress, however, that even this level of structuring is not trivial. Without facilitation, there is a risk that projects overstate their contribution or list a very long series of hoped-for impacts without prioritizing. One interviewee gives the example of improved school performance: it is reasonable to expect that changes in discipline and routine supported by a project might contribute to better grades, but not to assume that all progress is due to project activities alone. In their view, the role of external support – from universities, intermediaries or more experienced organizations – is to help projects formulate theory of change measurement that are ambitious but cautious, avoid over-attribution, and focus on outcomes that can be observed and documented with available tools.

Finally, specialists return to the importance of combining different types of evidence. They advocate for a basic mix of simple indicators, structured stories and partner feedback, rather than for reliance on a single method. In this configuration, frameworks

such as theories of change are not ends in themselves, but working tools: they help projects and supporters clarify what to pay attention to, organize the everyday evidence already being produced, and frame results for different audiences. Used in this way, formal frameworks become one element in a proportionate approach to measurement, rather than a separate layer of work added on top of overstretched practice.

6.6 Proportionate adaptations: “conversion recipes”

The analysis so far indicates that small community projects already generate a significant amount of information about their work and its effects, but that this information is not always legible to funders operating through formal CFPs. Specialists interviewed for this thesis argue that the problem is not the absence of evidence, but the lack of light, feasible structures that help “translate” everyday evidence into formats that funders can understand and compare. They also emphasize that such structures should respect the complexity of the context and the limited resources of small projects.

Across project and specialist interviews, certain ideas recur about how this translation might happen in practice. This section organizes these ideas into a set of proportionate adaptations – here described as “conversion recipes” – that build directly on existing routines and on the suggestions made by interviewees. They are not presented as a prescriptive toolkit or as full frameworks in the sense of large-scale evaluation systems. Rather, they summarize small bundles of tools that participants themselves point to as ways of making better use of the evidence already produced, and of helping funders recognize this evidence as valid.

6.6.1 A short project sheet grounded in everyday language

A first adaptation concerns how projects present themselves. Many coordinators already have a clear way of explaining, in conversation, what they do, who they serve and what changes they see. However, this narrative is often not captured in a stable written form.

Several project leaders describe situations in which they had to write similar information repeatedly for different CFPS or felt that they “knew what to say” orally but struggled to adapt to their language. In light of this, specialists suggest that a simple step would be to support projects to develop a short “project sheet” that condenses this explanation in two or three pages. This sheet would include: a brief description of the project’s origin and motivation; the main activities and their rhythm (for example, football training three afternoons a week, art workshops on Saturdays); the target group (age ranges,

approximate number of participants, main characteristics of the territory); and a concise description of what counts as a good result in their own terms (changes in discipline, confidence, school engagement, family relations, trajectories).

Such a document would mirror the kinds of narrative fields that CFPs already ask for, but would be written initially in the project's own language, possibly with the help of an intermediary or a university partner. It could then be adapted to different CFPs without having to be rewritten from scratch each time.

6.6.2 A three-indicator sheet aligned to existing practice

A second adaptation would focus on making explicit a small number of indicators that projects are already, in practice, observing or recording. Specialists suggest that, for small initiatives, one to three key outcome indicators are usually enough and more realistic than long indicator lists.

Based on the themes that appear most consistently in the interviews, such a "three-indicator sheet" might include, for example: one indicator related to participation and continuity (for example, the number or proportion of children who attend regularly over a school year); one indicator related to school engagement (such as the proportion of participants enrolled in school or reporting more regular attendance); and one indicator related to socio-emotional or behavioral change (for instance, the number of children whose behavior in class is described by teachers as more concentrated or respectful).

Each indicator would come with a simple definition and a note on how it will be tracked, using tools that the project already has: attendance lists, brief check-ins with families or teachers, and short observation notes. The aim is not to impose new, complex data collection routines, but to crystallize what is already being observed into a format that can be easily inserted into CFPs forms.

6.6.3 A very light baseline at intake

One recurring gap in the comparison with CFPs demands is the absence of formal baselines. Projects know, from conversations and observations, how children were before joining, but this is rarely written down in a structured way.

A proportionate adaptation mentioned by specialists would be to introduce a very light baseline procedure at intake, integrated into conversations that projects already have with children and caregivers. For example, at the first registration or meeting, a coordinator

could fill in a short form with a few key pieces of information: whether the child is enrolled in school; how frequently they have been attending; whether there are particular behavioral or emotional difficulties mentioned by parents or teachers; and what the child and caregiver expect from the project.

This baseline form could be kept in the same folder or digital file as the attendance list. Over time, when projects notice changes – a child who was not enrolled now is, a previously very shy participant begins to perform on stage, a parent reports fewer conflicts at home – they could add brief notes to the same file. Without being burdensome, this would create a simple before–after record that can be used in reports and to respond to CFPs questions about how change is measured.

6.6.4 Structured story prompts and trajectories

Stories of individual change are one of the most powerful forms of evidence that projects already have. Coordinators speak at length about particular children, adolescents and women whose trajectories illustrate the impact of sustained participation. Specialists argue that such narratives are not secondary but central to understanding complex social change.

A conversion recipe here would be to adopt a simple story template that helps to systematize these narratives. For instance, each story could be written with a few prompts: who is the participant (without full identification, but with basic characteristics); what was their situation when they arrived; what activities they took part in; what changed over time; and how the project contributed to this change. Adding approximate dates and linking the story to basic quantitative information (how long they participated, in which activities) would strengthen its evidential value.

Projects could aim to write a small number of these stories per year, selecting cases that represent different patterns (for example, a child with special needs, a teenager who becomes a monitor, a mother who finds more stability after joining a group). These narratives could then be attached to reports and applications, complementing indicators with concrete illustrations.

6.6.5 Simple partner confirmations

Another source of evidence already present in practice is the feedback that projects receive from schools, health professionals, social workers and local partners. Teachers report that children are more focused, psychologists notice emotional changes, and

community organizations acknowledge that activities help to “hold” children in protective spaces.

At present, much of this feedback circulates informally, through conversations and messages. One adaptation mentioned in interviews would be to encourage and support projects to ask partners for short written confirmations. These could be simple letters or emails stating the type of collaboration, what the partner has observed over time and how they see the project’s contribution to the community. For schools, for example, a note could mention that a group of students who attend a given project show improvements in attendance or behavior, even if these are not quantified precisely.

Such confirmations would respond directly to CFP requirements for attestations of technical capacity and partnership history, but would be grounded in existing relationships rather than generated solely for compliance.

6.6.6 An administrative “starter pack”

Finally, specialists and project leaders alike point to the need for support on the administrative side. While some requirements in large CFPs will remain beyond the reach of very small initiatives, an intermediate step is possible.

Several interviewees emphasize that having a minimum set of documents ready would already make a difference. An “administrative starter pack” could include a checklist of basic documents that projects can realistically obtain and maintain, such as: an up-to-date CNPJ certificate; a simple statute or founding document; bank account information; a basic cashbook or spreadsheet showing monthly income and expenses; and copies of any existing council registrations or partnership agreements. This pack would not replace the more demanding documentation required by some CFPs, but would put projects in a better position to apply to simpler ones and to negotiate with supporters.

Crucially, doing this work requires time and often some external accompaniment. Specialist interviews highlight the role that intermediaries, networks and universities can play in helping small projects assemble and update such packs, or in pooling administrative support across several initiatives so that no single coordinator has to carry the entire burden.

Taken together, these conversion recipes are not a blueprint, but a set of feasible shifts that emerge from the material. They start from what projects already know and do, and

show how this could be turned into something more legible without turning everyday practice into a bureaucratic exercise. Read in this way, they indicate concrete ways in which the triangle might be brought closer to balance. The extent to which such adaptations can and should be taken up in practice is discussed more directly in the following chapter.

6.7 Synthesis back to purpose–feasibility–legibility

The results presented in this chapter answer the three research questions in an integrated way. They show how small community projects understand and register impact in everyday practice (RQ1), what CFPs and funders demand in terms of indicators, documentation and monitoring (RQ2), and where these practices and demands align or misalign (RQ3). The purpose–feasibility–legibility triangle introduced earlier offers a simple lens to synthesize these findings and to outline possible ways forward.

On the purpose dimension, projects and funders partly converge and partly diverge. Project leaders speak of impact in terms of concrete, situated changes: children arriving on time and staying through the year; adolescents gaining confidence to perform or to teach; families experiencing fewer conflicts and more support; women leaving violent relationships and finding spaces to rebuild routines. Their evidence practices serve multiple purposes at once: they help to follow participants' trajectories, to adjust activities, to feed back to families and schools, and, when needed, to show supporters that the work is worth continuing.

Funders, especially those operating through large CFPs, also seek to promote social and environmental change and to reach vulnerable groups. However, the primary purpose of their measurement demands is often accountability and comparison across a portfolio of projects. Indicators, workplans and monitoring templates are designed to allow selection committees and internal teams to assess consistency, take decisions and report upwards within their own organizations. This difference in emphasis explains some of the tension observed in the interviews: project leaders feel that they know very well what matters and what is changing, but struggle to see how this maps onto the categories that funders consider valid.

On the feasibility dimension, the results point to a persistent imbalance. Projects in this study operate with small teams, sometimes only one or two central coordinators supported by volunteers. They work in territories where public services are intermittent, spaces are

shared, and resources must be stretched to cover basic activities. Within these constraints, they have nonetheless developed a dense fabric of everyday evidence: attendance lists, intake conversations, check-ins with families, stories and photos, archives of events and recognitions. These practices are feasible because they are embedded in routines and respond to immediate needs.

By contrast, many CFP demands presume a different organizational profile: one with stable staff dedicated to administration and evaluation, access to accountants and lawyers, the ability to produce detailed budgets and reports, and time to learn and navigate online platforms. For the smallest projects, meeting these demands would require a reallocation of scarce time away from core activities or external support that they often do not have. Specialists interviewed here underline that, when requirements overshoot feasibility, they not only exclude many initiatives from funding opportunities but also risk distorting practice, encouraging projects to mimic forms without necessarily improving their own learning.

The legibility dimension captures how evidence travels across these worlds. Locally, projects are highly legible to participants, families and nearby institutions. Neighbors see children going to and from activities; parents notice changes in behavior; teachers perceive improvements in concentration; local organizations and companies observe the presence and persistence of projects over years. Everyday evidence, while informal, is persuasive in these relationships.

At the level of formal funding, however, this evidence often loses visibility. Selection committees do not see daily routines; they read written proposals, scoring sheets and summaries. In that arena, what becomes legible is not only what projects do, but how they can present what they do: their capacity to write coherent justifications, to fill indicator tables, to assemble administrative documents and to show a formal history of partnerships and financial management. Projects that are deeply embedded in their communities but have little support in translating their practice into this format find themselves at a disadvantage.

The fit–gap analysis suggests that misalignment is not evenly distributed across dimensions. In terms of substantive focus and basic counts, there is a strong fit: projects work in priority areas and can show how many people they reach. In terms of monitoring routines, there is enough alignment to build on: attendance lists and stories can form the

basis of simple indicators and narratives. The most structural misfits lie in legal–administrative requirements, in the sophistication of financial controls expected, and in the written form and language of proposals.

The proportionate adaptations sketched in Section 6.6 summarise directions that emerged in the material for bridging some of these gaps. They are not a call for small projects to adopt complex frameworks or to shift their primary accountability away from participants and communities. Instead, they draw together modest changes mentioned by projects and specialists – a project sheet, a three-indicator summary, a light baseline, structured stories, partner confirmations and an administrative starter pack – that could, in principle, help make existing evidence more visible to funders without overburdening teams.

At the same time, the results indicate that the responsibility for improving fit does not lie solely with projects. Funders and intermediaries have significant room to adjust their own practices: by simplifying documentation requirements for smaller grants; by recognizing everyday evidence as valid when structured proportionately; by providing feedback to unsuccessful applicants; and by investing in shared support structures that help community initiatives navigate legal and administrative demands. Specialists emphasize that such adjustments would not weaken accountability. On the contrary, they can make it more realistic and thus more likely to be met in practice.

The empirical material assembled in this chapter shows that small community-based projects in Rio de Janeiro and surrounding areas are neither “empty” of evidence nor indifferent to impact. They have developed pragmatic ways of observing and documenting change that suit their purposes and constraints. The challenge lies in ensuring that these ways of knowing are not lost in translation when they encounter institutional systems designed with very different organizations in mind. The purpose–feasibility–legibility triangle provides a simple heuristic way to think about this challenge.

7. Discussion

7.1 Bringing the threads together

This chapter brings together the empirical material and interprets it through the theoretical framework. It returns to the three research questions that motivate the study: how small, community-based projects in Brazil define and evidence “impact” with the resources they

have (RQ1), what funders and CFPs require and how project leaders experience these demands (RQ2) and where everyday evidence practices and formal requirements align or misalign (RQ3). What counts as credible evidence depends on what the evidence is for, what can realistically be sustained over time, and what must be legible to external audiences. The purpose–feasibility–legibility triangle introduced earlier provides a structured lens for analyzing these tensions without reducing them to a simple distinction between "good" and "bad" evaluation.

The discussion develops three contributions. First, the cases show that small projects often do measure and document change, but they do so through practices embedded in relationships and everyday routines. This makes their evidence difficult to compress into standardized and aggregated reporting formats. Second, the analysis indicates that documentation and monitoring demands shape not only post-grant reporting but also access to funding. Some initiatives exclude themselves out of opportunities when requirements appear disproportionate to their capacity, so "fit" influences access before any reporting begins. Third, intermediaries function as practical support for measurement. They help projects translate the evidence they already produce into materials that align with funder expectations. In doing so, they reduce the burden of "figuring out the format" and make reporting more feasible for capacity-constrained teams.

A key interpretative claim running through the chapter is that many tensions around impact evidence are structural rather than motivational. On the project side, evidence is often designed for proximity to the community. It supports follow-up, coordination with families and local partners, and small adjustments over time. On the funder side, evidence is often designed to be easy to compare and report across many projects. It supports comparability across grantees, procedural accountability, and organizational risk management. These orientations are not inherently incompatible, but they privilege different forms of proof and different ways of documenting.

The empirical material shows multiple strategies for bridging these arenas, including adapting local documentation to required categories, relying on intermediaries to translate, or avoiding CPFs perceived as unrealistic. Across cases, the main constraint is rarely a lack of willingness to show results. More often, it is the fit between (i) the purpose evidence serves in the project, (ii) what is feasible to sustain as a routine, and (iii) what is legible to external audiences in the formats demanded by funding mechanisms. This is

why the purpose–feasibility–legibility triangle is useful. It makes visible how evidence demands can escalate independently of feasibility, and how legibility pressures can exclude learning-oriented documentation even where learning would be more decision-relevant for implementers.

7.2 Everyday evidence and local definitions of impact (RQ1)

The first research question asked how small, community-based projects define "impact" and how they make it visible in everyday practice. In these cases, impact is not introduced through an indicator framework. It is explained through people's trajectories and the kinds of change that matter to their daily life. In none of the cases, project leaders started by listing indicators. Instead, they start with what it is that is shifting for participants and families, and with the signs that they see was proof that the project is "working" in the territory. This is close to what Brazilian authors describe as "everyday evidence," where evaluation is tied to lived realities and practical purposes rather than treated as a purely technical reporting exercise (Lacerda, 2014; Silveira, 2020; Monteiro, 2010). When the project leaders give examples, they point to concrete changes that are easy to recognize locally, such as a child who used to struggle to focus, now completing activities or adolescents learning to take on more responsibilities over time.

Their way of defining impact also helps explain why many of the outcomes that the project leaders emphasize do not translate easily into standardized metrics. This speaks into the capability lens frame of value, in terms of what people are effectively able to do and to be in everyday life (Nussbaum, 2011). Many cases fit this register, as their main focus is on participation, agency, persistence, and the ability to keep routines. A relational lens adds another layer: a big part of what projects "produce" sits between people, in trust, belonging, and cooperation (Donati, 2014). These are often visible to those who live and work in the community, even though they are only partially captured by formal indicators. In that sense, the cases support the argument that impact for small projects is not easily reduced to service counts or single composite scores (Vanclay, 2003; Vanclay et al., 2015).

In practical terms, the projects do more monitoring than a first glance at their spreadsheets might suggest. Attendance lists, registration records, WhatsApp groups, short case notes, and recurring conversations with families and schools form a running picture of participation and change over time. This material is not always systematized in the way external manuals recommend, but it is actively used. Coordinators rely on it to notice who

is drifting away, decide when follow-up is needed, adjust schedules, respond to conflicts, and maintain credibility with local partners. In that sense, monitoring is less a separate "evaluation activity" and more part of how the work is managed and learned from in real time (Monteiro, 2010; Silveira, 2020). Many project leaders value monitoring for internal use, rather than for external purposes.

At the same time, the findings complicate the common portrayal of small organizations as having "no data." The projects often do have information, but it is distributed across people, routines, and informal records. Leaders can reconstruct long trajectories from experience and basic documentation, even when they cannot easily produce aggregated summaries across cohorts. The practical difficulty is not a total absence of evidence. It is the translation of locally meaningful knowledge into portable formats that match what funders and selection committees tend to value.

Finally, the results speak directly to proportionality. Authors such as Prates Rodrigues (2011, 2018, 2021) argue that evaluation should be scaled to real conditions and to the decisions at stake. The cases show what this looks like from the inside. Small teams prioritize routines they can actually maintain and that respond to local expectations, especially from families and schools. This often means that more demanding designs remain out of reach, including long-term follow-up or counterfactual evaluation. Feasibility here is not only technical capacity. It is also about what can be sustained without pulling time away from delivery work and from the relationships that make these projects possible (Gertler et al., 2016).

7.3 Measurement demands, CFPs and institutional logics (RQ2)

The second research question examined what CFPs and funders require in terms of indicators, documentation, and monitoring, and how project leaders experience these demands. Across the material, these requirements read less like a neutral "request for information" and more like an accountability infrastructure. They specify not only *what* should be reported, but also *what a proper project is supposed to look like* and how it should present itself to outsiders. In this sense, measurement demands operate as governance tools that stabilize comparability, auditability, and organizational risk management across a portfolio (DiMaggio & Powell, 1983; Meyer & Rowan, 1977; Greenwood et al., 2011).

This is consistent with the CSR discussion seen in chapter 4.3, where auditability and standardization become attractive because they produce evidence that can be reviewed, compared, and defended to external audiences (Banerjee, 2008; Barnett et al., 2020). In the CFP context, this has consequences beyond reporting. It also shapes which projects become visible and competitive in the first place. The result analysis suggests that CFPs tend to reward applicants who already know how to translate their work into the expected formats and narratives. A second mechanism that reinforces this dynamic is the lack of feedback to unsuccessful applicants. In the interviews, leaders describe being told that they "almost passed" without knowing what was missing, which makes it difficult to learn the rules of legibility across cycles. Over time, this can contribute to the drop-off observed in the cases: some initiatives stop applying, not because they do not have results to show, but because they cannot see how to translate those results into what selection committees recognize as "good enough" evidence.

Across the analyzed CFPs (Appendix A), there is a familiar core package. Applicants are typically expected to present a problem diagnosis, objectives, target group, planned activities, indicators, baselines and targets, and a monitoring plan, combined with legal and financial documentation. What matters here is not only the content, but the assumptions behind it. Many CFPs implicitly presuppose administrative routines, documentary traceability, and enough staff time to maintain platforms, upload bundles of documents, and keep reporting cycles on schedule. The heaviest requirements in the sample concentrate on larger public and corporate calls embedded in formal compliance systems, where procedural diligence to boards, regulators, and internal controls is a central concern. This pattern aligns with how standardized reporting infrastructures stabilize expectations by creating templates that can be reviewed, compared, and defended, even when initiatives and contexts differ (Global Reporting Initiative, 2021; Global Impact Investing Network, 2019).

From the perspective of project leaders, these demands are often experienced inconsistently. Several acknowledge that workplans, matrices, and indicator tables can help sharpen objectives and make activities easier to explain. At the same time, many describe the process as disproportionate to their reality. The burden is not only the indicator logic itself, but the wider scaffolding that comes with it: repeated uploads of similar documents, tight deadlines, form-driven language that does not match how projects talk about change, and limited support for converting everyday practice into

institutional categories. Brazilian practitioner-oriented literature similarly notes that evaluation and monitoring demands can overshoot the routines and infrastructures available to smaller organizations, particularly where information systems are dispersed and consolidation is difficult (Prates Rodrigues, 2018; Monteiro, 2010).

A second pattern concerns the gap between compliance-oriented reporting and learning-oriented use. Several leaders describe completing indicator fields and logical frameworks primarily because “the call asks for it”, often choosing what is easiest to count or easiest to justify rather than what actually guides practice. This reflects a well-established institutional mechanism. Under legitimacy pressure, organizations may adopt formal structures that are culturally valued and reviewable, even when these structures remain loosely connected to day-to-day work (Meyer & Rowan, 1977). Kah and Akenroye (2020) similarly show that tool choice in impact measurement is often shaped by external expectations for legitimacy and accountability, not only by internal learning needs. From this perspective, “minimal compliance” is not simply a quality problem. It can be a rational organizational response when requirements exceed feasibility and when the primary payoff is external acceptability rather than internal decision use. Greenwood et al. (2011) further help explain why this can persist under institutional complexity, as organizations compartmentalize competing logics, stabilizing outward-facing routines while relying on different internal practices (Greenwood et al., 2011).

Finally, the findings suggest that measurement demands can shape who applies for funding in the first place. The interviews showed that some leaders decide not to apply because they already anticipate that documentation, co-financing, and monitoring requirements will not fit their reality. Few project leaders rely on intermediaries who screen opportunities and absorb part of the administrative burden. This implies that CFPs do not only shape reporting after grants are awarded. They also shape who enters formal funding portfolios in the first place. As a result, the set of initiatives that funders “see” may be filtered by legibility and administrative fit, not only by substantive social value.

Taken together, these patterns set up the next section’s focus on fit and misfit (RQ3). The same evidence demand can be reasonable from a portfolio-governance perspective (legibility) while being unworkable from an implementation perspective (feasibility) or misaligned with what local actors treat as meaningful change (purpose). This is why the purpose–feasibility–legibility triangle is analytically useful. It makes it visible how

demands for standardized, auditable evidence can escalate independently of feasibility, and why compliance-oriented documentation may crowd out learning-oriented routines in small, community-based settings.

7.4 Negotiating fit: purpose, feasibility and legibility in practice (RQ3)

The third research question asked where everyday evidence practices and CFP requirements align or misalign, as well as what feasible adaptations could make meaningful evidence more legible. The central argument is that "fit" is not only about choosing the right tool. It depends on why evidence is being produced, what can realistically be sustained, and what counts as credible in the formats used by funders. These three aspects often pull in different directions, even when everyone agrees that evidence should be "good". Therefore, misfit is better understood as a recurring, structural issue, not as a lack of motivation or competence on the project side.

Looking first at purpose, projects and funders overlap in some areas, but the priorities still diverge in a consistent way. Project leaders mainly mobilize evidence to support participants, adjust activities and communicate with families and local partners. Funders, in turn, rely on evidence for selection decisions, compliance routines, accountability to boards or regulators, and portfolio-level reporting. The same artefact can therefore serve different purposes. An attendance list helps a coordinator notice who is drifting away and decide on follow-up, while for a selection committee it functions primarily as proof of delivery and traceability. This is consistent with the argument that measurement serves multiple purposes and that methods should be assessed in relation to the decisions they are meant to inform (Ebrahim & Rangan, 2014; Kah & Akenroye, 2020). It also clarifies why misunderstandings are common.

Turning to feasibility, it repeatedly emerges as the main constraint across cases. Small projects often depend on one or two central figures, operate with volatile funding, and have limited access to administrative systems and dedicated evaluation staff. Within these limits, they build routines that are sustainable and decision-relevant in daily work. Many CFP requirements, by contrast, implicitly assume organizations with separated roles, stable digital infrastructures, and staff time for reporting cycles. When these assumptions meet grassroots realities, feasibility breaks down in predictable ways. Requirements are not met, are met only partially, or are met at the cost of stretching already thin capacity and diverting effort away from delivery and relationships. Brazilian practitioner-oriented

critiques of disproportional evaluation demands, and “fantasy” impact claims help interpret this as a credibility problem, not simply a capacity shortfall (Prates Rodrigues, 2018; Prates Rodrigues, 2021; Lacerda, 2014).

Legibility concerns to whom evidence “makes sense” and in what format. In local settings, projects can be highly legible. Residents observe long-term presence, schools notice behavioral and routine changes, and families experience shifts in everyday stability. In this arena, narratives, reputational cues, and relational knowledge carry weight because they circulate through ongoing relationships. In the CFPs arena, legibility is mediated by forms, scoring rubrics, and document bundles, and evidence must be translated into standard categories to be recognized as credible. A key implication is that legibility is unevenly distributed. Projects with similar practices can appear very different to external reviewers depending on their access to administrative support, writing capacity, and intermediary assistance.

The analysis also identifies spaces where fit is possible. Specialists and intermediaries describe “conversion” strategies that start from what projects already do, such as attendance tracking, informal records, and brief narratives, and stabilize this into small evidence packages that travel more easily. Rather than importing a full framework such as SROI, these strategies emphasize selective tools that are proportionate to capacity, including a concise Theory of Change, a small set of negotiated indicators, and structured change narratives that preserve context while remaining reviewable. This aligns with Social Impact Assessment traditions that emphasize participation and attention to context when interpreting significance (Vanclay, 2003; Vanclay et al., 2015). The empirical material suggests that the building blocks for this kind of conversion already exist in practice. The challenge is to recognize and stabilize them so that feasible, purpose-driven routines become legible to external audiences, without being displaced by compliance-only documentation.

7.5 Implications for practice

Taken together, the findings suggest that improving impact evidence is less about asking small initiatives to “measure more,” and more about redesigning evidence expectations so that purpose, feasibility, and legibility stay aligned. The implications below focus on how community-based projects, funders and CFP designers, intermediaries, and the

broader field can reduce the gap between evidence-in-practice and institutionally legible proof without turning measurement into a parallel burden.

For community-based projects, the main implication is to treat everyday documentation as the starting point for a small but coherent evidence routine. What can look “informal” from the outside, such as attendance lists, basic registries, WhatsApp histories, short case notes, and recurring conversations with families and schools, can be organized into an evidence portfolio that remains feasible. The practical move here is modest systematization. Teams can agree on a shared way to record a small set of signals they already use to steer work, for example entry, continuation, drop-out, referral, and re-engagement, and then pair these signals with a small number of short trajectories that make mechanisms of change visible. This strengthens internal learning and coordination while also making external communication easier. At the same time, the thesis cautions against adopting heavy infrastructures, such as dashboards with frequent reporting cycles, complex scoring tools, or monetization frameworks, when the decision value is unclear and sustaining the system would compete with delivery and relationships (Ebrahim & Rangan, 2014; Prates Rodrigues, 2018; Prates Rodrigues, 2021).

For funders and CFP designers, the findings imply that proportionality has to be built into them, not left to grantees to absorb. If a program claims to target grassroots initiatives, the call should explicitly assume low administrative capacity and reflect that assumption in its structure. In practice, this means reducing redundant documentation, clarifying why each requirement exists, and aligning evidence demands with what is realistically observable in community settings. It also means closing the loop for applicants who are not selected. In the empirical material, the absence of feedback comes up as a concrete barrier: leaders describe not knowing what they did wrong and therefore not knowing how to improve, which contributes to some projects stepping away from CFPs altogether. Even minimal, standardized feedback would make the criteria more learnable and can reduce self-exclusion, especially for initiatives that do not have access to grant-writing support. Accountability is not abandoned, but reframed toward material and reviewable evidence rather than exhaustive reporting for its own sake (Global Reporting Initiative, 2021). A practical way to do this is tiering. Minimum evidence requirements can apply to small grants and early-stage initiatives, with escalation only when grant size, organizational maturity, and risk justify it. It also means accepting more than one valid format. Allowing a combination of a small indicator set and structured narratives can

support both legibility and context, without forcing projects into long pre-defined indicator lists that do not fit how change happens locally (Global Reporting Initiative, 2021; Feor et al., 2023).

For intermediaries, the implication is that translation is not a peripheral service but a form of measurement infrastructure. Intermediaries reduce legibility costs by helping projects convert embedded documentation into call-ready formats, align local narratives with institutional language, and stabilize routines that can be sustained over time. This matters because it can prevent compliance from becoming a parallel system that is disconnected from delivery (Prates Rodrigues, 2018). It also has an equity dimension. Translation capacity reduces the advantage of having exceptional administrative and writing resources. Strengthening this function through dedicated funding for technical assistance, shared templates co-designed with projects, collective training spaces, and helpdesk-style support during application windows may be more effective than expecting each initiative to build a full evaluation infrastructure alone.

For the broader field of impact measurement, the thesis suggests shifting the starting point. Tools for small initiatives should begin with existing evidence practices rather than with an idealized model of formal evaluation. Embedded, relational monitoring should not automatically be treated as an obstacle to "proper" evaluation. It can be treated as a resource that can be structured and made communicable. From this perspective, proportionate approaches are not simply scaled-down versions of large-program methods. They are designs that preserve learning and relevance while improving portability. This aligns with guidance that distinguishes different purposes of measurement and warns against default escalation toward more demanding methods as a proxy for quality (So & Staskevicius, 2015; Rogers, 2014). The central challenge is to avoid evaluation "fantasy," meaning claims and requirements that exceed what data, context, and capacity can credibly sustain, while still enabling accountability that is meaningful and fair (Prates Rodrigues, 2021).

7.6 Translation work, legibility labor, and the hidden infrastructure of impact reporting

A recurring theme across the findings is that the hardest part of "impact measurement" is often not measurement in the narrow sense. It is translation. The projects in this study are not working with a blank slate. They already observe change, notice patterns, and keep

traces that allow them to follow participants over time. What becomes difficult is turning that embedded knowledge into forms that travel across institutional systems. That conversion takes time, specific skills, and, in many cases, external support.

This is especially visible at the application stage. CFPs do not only ask projects to describe what they do. They also require them to speak a particular language. Projects are expected to present themselves as structured plans with predefined objectives, indicator choices, and monitoring routines, alongside administrative proofs that make the organization legible and auditable. From the perspective of portfolio governance, this is understandable. It helps funders compare applicants and justify decisions. From the perspective of small initiatives, however, it means that "impact evidence" becomes a gatekeeping format. A project can be doing strong work and still appear weak if it cannot package that work into the expected categories.

The empirical material suggests that legibility is produced through a series of small steps that are easy to miss when one focuses only on the final application. Leaders interpret CFPs and try to anticipate what reviewers will recognize. They rewrite local aims into institutional keywords. They decide what can be framed as a "result" without overpromising. They assemble documents that demonstrate organizational capacity, even when much of the work is carried by informal routines and relationships. This work is not purely technical. It involves judgment about what to foreground, what to simplify, and what to leave out. It also often happens under time pressure, with little room for iteration.

This matters because the labor of legibility is unevenly distributed. Two initiatives can have similar evidence practices on the ground and still face very different odds in competitive CFPs, depending on whether they have administrative help, writing confidence, digital tools, and intermediary support. In that sense, legibility is not only a communication issue. It becomes an equity issue. It also suggests a broader definition of "capacity." Capacity is not only the ability to track outcomes. It is also the ability to translate outcomes into institutional formats, comply with documentation regimes, and perform competence in a genre that is often closer to professional grant writing than to community work.

Institutional theory helps interpret why this translation pressure is so persistent. When funders and evaluators operate under expectations of comparability and accountability, they tend to reward applications that resemble each other in structure and language. Over

time, recognizable templates become taken-for-granted markers of seriousness. This does not require bad faith. It is a predictable dynamic in fields where legitimacy is distributed through formalized procedures and where evaluation must be conducted at a distance. From the perspective of small initiatives, the effect can still be discouraging. When the "right" format becomes a condition of visibility, projects learn that what counts is not only what they do, but how convincingly they can describe it in the expected register.

This also clarifies why over-demanding measurement can generate performative responses. If CFPs ask for extensive indicator systems, baselines, targets, and detailed monitoring plans, but do not account for feasibility, applicants may feel pushed toward optimistic claims or overly tidy models of change. The risk is not simply that projects "do not measure." The risk is that they learn to write as if they measure in ways that cannot be sustained. In that situation, measurement becomes a parallel system, convincing on paper but weakly connected to delivery. Legibility then wins over learning. Evidence becomes something produced mainly for selection and reporting cycles, rather than something that strengthens practice.

The documentary analysis reinforces this interpretation. Many CFPs embed a particular imagination of organizational life, with separated roles, stable record-keeping, predictable timelines, and reporting routines that can be maintained regardless of volatility on the ground. The more a call assumes this infrastructure, the more "fit" depends on administrative conditions rather than on the substantive quality of community work. This helps explain why some leaders in the study describe self-exclusion as a rational response. They do not necessarily doubt the value of their work. They doubt their ability to compete in a documentation-heavy arena without sacrificing scarce time and attention.

Intermediaries appear in the findings as a key response to this translation problem. In practice, intermediary Organizations do more than "help." They function as infrastructure for measurement and legibility. They reduce interpretation costs by explaining CFPs and reviewer expectations. They stabilize routines by offering templates and checklists that projects can realistically maintain. In some cases, they act as editors, converting embedded documentation into application-ready narratives. This can prevent the worst form of decoupling, where reporting becomes entirely separate from work, because it supports a more faithful representation of practice. At the same time, this intermediary role raises its own questions. If access to translation support determines access to funding,

then intermediaries become part of how inequality is reproduced or reduced in the field. Translation capacity becomes a scarce resource that can concentrate opportunity among those who are already better connected.

This is also where the thesis intersects with debates about standardization and commensuration. The push for auditability and benchmarking does not necessarily follow from a settled agreement on what impact is, or how it should be measured. It can intensify even when standards remain contested and when important dimensions of social change resist reduction to comparable units. In practice, this can narrow what is recognized as value. Relationship-building, stability, belonging, and expanded agency may remain central to participants' lived change, yet travel poorly when a call expects short, quantified signals. The concern is not that these outcomes disappear in reality. The concern is that they disappear from what is rewarded.

A further implication concerns the role of AI tools, which is relevant here because AI is increasingly available as low-cost support for writing and summarization. On one hand, AI can help projects organize applications, clarify language, and reduce the burden of producing legible documents. That could lower dependence on expensive grant-writing expertise. On the other hand, AI can also intensify standardization. It can smooth local specificity into generic development language, and it can inflate claims by producing confident-sounding narratives that exceed the underlying evidence. For grassroots initiatives navigating high-stakes accountability, this creates a delicate trade-off. AI can support legibility, but only if it is used cautiously, with strong author oversight and a clear commitment to contribution-oriented claims rather than inflated attribution.

Taken together, these reflections suggest that improving impact measurement for small initiatives requires taking translation seriously as part of the core problem. The challenge is not only to design better indicators or lighter reporting systems. It is to redesign evidence environments so that they recognize the kinds of knowledge that exist in community practice, and so that access to funding does not depend on hidden capacities that are unevenly distributed. Practically, this points toward shared translation infrastructure, such as helpdesk-style support during application windows, co-designed templates, and tiered requirements that scale with grant size and risk. Ethically, it supports treating proportionality as a fairness criterion. If a call claims to target grassroots initiatives, then the burden of legibility should not be outsourced to those initiatives alone.

In that sense, one of the most important questions raised by this thesis is not whether grassroots projects can "do impact measurement" in the abstract. They already produce evidence in everyday practice. The more important question is what kinds of evidence arrangements allow that practice to be recognized without forcing it to mimic the evaluation infrastructure of large organizations. When translation is acknowledged as labor and as infrastructure, the path forward becomes clearer. It lies in building bridges between everyday evidence and institutional legibility, without letting those bridges become another barrier.

8. Conclusion

8.1 Revisiting the research questions

This thesis started from a complex and recurrent problem. Small, community-based projects in Brazilian urban peripheries are increasingly asked to "show impact," yet many of the tools and templates circulating in policy, philanthropy, and impact funding assume levels of data, time, and infrastructure that these projects do not have. The study addressed this problem through interviews with project leaders and specialists, alongside a documentary analysis of calls for proposals. It asked three questions: how small projects define and evidence impact in everyday practice (RQ1); what funders formally ask of them (RQ2); and where these practices and demands align or misalign, and with what consequences (RQ3).

In response to RQ1, the empirical material shows that small projects do produce and use evidence, but on terms that begin with people, trajectories, and relationships rather than with indicator frameworks. Leaders describe impact in concrete, situated ways, including changes in participation, routines, confidence, and everyday stability, and they monitor these changes through pragmatic documentation that helps teams notice who is engaging, who is drifting away, and what kinds of change appear significant over time. These routines are embedded in daily work and oriented first to local public, including participants, families, nearby schools, and community partners, rather than to distant funders.

In response to RQ2, the CFP corpus and specialist interviews reveal a different center of gravity. CFPs and program guidelines emphasize legal and financial documentation, formal governance, detailed workplans, indicator tables, monitoring plans and platforms, and audit-ready records. These requirements make applicants comparable and help

funding institutions justify decisions and report upwards. In this logic, indicators and templates function not only as learning tools but also as instruments of selection, portfolio management, and organizational risk control. Even CFPs that explicitly target small or community-based initiatives frequently reuse this language, with only partial adjustment to differences in scale and administrative capacity.

In response to RQ3, the evidence shows that projects adopt a range of strategies at the boundary between local practice and institutional demand. Some adapt existing documentation to fit CFP categories. Some rely on intermediaries to translate local realities into standardized formats. Others avoid CFPs whose requirements they experience as incompatible with their size or way of working. Across cases, the main constraint is rarely a lack of interest in showing results. More often, it is the fit between purpose, meaning what information is needed and by whom, feasibility, meaning what can realistically be collected and sustained as a routine, and legibility, meaning what counts as credible evidence to specific audiences.

The purpose–feasibility–legibility triangle proposed in this thesis offers a way to describe this dynamics without defaulting to a deficit story. Projects are not simply failing to evaluate. They are navigating multiple expectations with thin resources while trying to keep their priorities, participants, relationships, and continuity, at the center of their work. Funders and CFP designers are also not simply indifferent to context. They operate under strong pressures for comparability and accountability. The central issue is how these pressures meet in concrete arrangements, and how far “fit” can be negotiated through proportional expectations and translation support.

8.2 Main findings and contributions

This thesis makes three interconnected contributions: an empirical contribution on how evidence is produced, used, and demanded in small community-based projects, an analytical contribution in the form of the purpose–feasibility–legibility lens, and a practical contribution that translates these insights into proportionate design principles for key actors in this field. Taken together, these contributions support the thesis proposition that impact measurement becomes more defensible and more useful when it is designed for fit, rather than imported as a standard package assumed to work across contexts and organizational forms.

Empirically, the study shows that “everyday evidence” in small projects is denser and more intentional than it often appears from the outside. Teams routinely observe and interpret change through attendance and participation patterns, that function as operational tools that support coordination, follow-up, and learning. In this sense, the thesis contributes a grounded account of how evidence is produced within practice, through proximity to participants and continuous interaction with local partners.

At the same time, the thesis documents how these forms of evidence sit uneasily with formal measurement demands in CFPs. Across the documentary material and specialist interviews, CFPs typically ask applicants to present clear objectives, predefined indicators, baselines and or targets, and monitoring plans, alongside extensive legal, financial, and administrative documentation and audit-ready records. These requirements serve informational purposes, but they also serve governance purposes. They make applicants comparable, support justification of funding decisions, and reduce organizational risk through traceability. The thesis therefore clarifies that impact measurement in this field is not only about learning what works. It is also a mechanism through which access is managed and legitimacy is granted.

A further empirical contribution concerns how projects respond at the boundary between local practice and institutional demand. Rather than a single pattern of compliance or resistance, the study identifies a range of coping strategies. These include adapting existing documentation into required categories, relying on intermediaries to translate practice into funder-ready formats, selectively complying by reporting what is easiest to count, or self-excluding from opportunities perceived as disproportionate. These strategies point to an upstream effect of measurement demands. Requirements shape not only reporting after funding is granted, but also participation and access before an application is even submitted. Documentation and monitoring requirements can therefore function as filters that influence which initiatives become visible in formal funding portfolios.

Analytically, the purpose–feasibility–legibility lens contributes by offering a concise way to hold these dynamics together. It clarifies that “fit” is rarely a purely technical issue. Evidence practices are shaped by what evidence is for (purpose), what can realistically be sustained as a routine (feasibility), and what formats and categories are recognised as credible by relevant audiences (legibility). Applied to the cases, the lens shows that misfit

often arises because these dimensions do not move together. Legibility demands can escalate independently of feasibility, and the purposes of evidence in everyday work can differ from the purposes of evidence in funding and reporting systems. This framing avoids reducing misalignment to a deficit story about "lack of capacity," and instead treats it as a structured tension produced by how accountability is organized.

Practically, the thesis points toward proportionate ways of designing evidence arrangements that reduce friction without turning measurement into a parallel burden. For projects, the implication is not to build heavy dashboards or complex measurement systems, but to stabilize what already exists into a modest and coherent evidence portfolio. This includes a concise articulation of intended change, a small set of locally meaningful indicators, and brief, systematically collected trajectories or narratives that preserve context while remaining communicable. For funders and CFPs designers, the implication is to build proportionality into the CFPs from the outset, through fewer and clearer requirements, evidence expectations that match the capacity profiles of targeted applicants, and acceptance of mixed evidence packages rather than one-size-fits-all templates. For intermediaries, the implication is that translation is not an optional add-on. It is a form of infrastructure that reduces legibility barriers and prevents compliance routines from becoming disconnected from delivery.

In conclusion, this thesis does not propose a new flagship method. It argues for fit-oriented measurement designs that keep purpose, feasibility, and legibility aligned, so that evidence remains decision-useful for implementers, credible to external audiences, and proportionate to the realities of small community-based work.

8.3 Limitations and directions for further research

This study has several limitations that matter for how its findings should be interpreted and for how far they may travel beyond the specific cases examined. These limitations also point to directions for further research that could strengthen, extend, or challenge the patterns identified here.

First, the empirical base is intentionally narrow and context-specific. The project cases are drawn from two communities in Rio de Janeiro and cover a limited number of initiatives. This design supported depth and a close reading of everyday practice, but it also limits the range of organizational forms, funding histories, and territorial conditions represented. The conclusions are therefore analytical rather than statistical. The thesis

identifies plausible mechanisms and patterns, but it does not claim to describe the Brazilian third sector as a whole. Future research could extend the analysis across regions, including other urban peripheries and rural contexts, and include initiatives with different levels of professionalization, legal status, and administrative capacity. Such comparisons would help clarify which misfit dynamics are broadly structural and which depend on local institutional arrangements.

Second, the study relies primarily on the perspectives of leaders and coordinators when describing project practices, priorities, and evidence routines. This focus fits the research questions, especially those concerned with how projects define impact and how they navigate external demands, but it leaves some angles under-explored. Participants, families, and front-line staff may interpret "impact," "good evidence," and the burden or value of monitoring differently. They may also experience documentation routines in ways leaders do not fully see, particularly where data collection shapes interactions, trust, or participation. Future research that integrates participant and family perspectives could deepen understanding of what kinds of evidence are considered meaningful locally and how evidence practices affect relationships and engagement.

Third, funder expectations were accessed mainly through CFPs, program guidelines, and a small number of specialist interviews. This provides a clear view of formal requirements, but it does not capture how evidence is interpreted, negotiated, and weighed inside funding institutions. The study did not observe internal deliberations, examine how selection committees interpret different evidence formats, or follow how monitoring requirements are implemented and enforced during grant periods. These processes matter because they shape whether a requirement operates as a strict threshold, a flexible expectation, or a symbolic checklist. Further research could examine decision-making more directly by including funder staff, committee members, and compliance teams, and by tracing how evidence travels through selection, contracting, monitoring, and renewal phases.

Fourth, while the study collected routine artefacts and documentary material alongside interviews, it did not attempt to reconstruct full quantitative series or reanalyze project data statistically. The analysis is qualitative and interpretive, centered on how evidence is produced, used, and translated. This is a strength for understanding meaning and practice, but it leaves open questions about what might be learned even from lightweight

quantitative patterns in everyday records. A complementary line of work could explore how far simple indicators derived from existing documentation, such as attendance trajectories or re-engagement patterns, can be systematized without imposing substantial additional burden. This would also help clarify when such indicators provide useful signals for decision-making and when they risk oversimplifying complex change processes.

Fifth, the study is cross-sectional. It captures a moment in which impact language and measurement demands are consolidating in Brazil, and in which projects, intermediaries, and funders are adapting to one another. A cross-sectional design cannot show whether evidence routines become more stable over time, whether "conversion" strategies are sustained or abandoned, or how relationships between funders and projects evolve as expectations rise or resources shift. Longitudinal research could follow specific initiatives, programs, or funding relationships over time to observe how measurement routines are negotiated, how trade-offs between learning and compliance play out, and what conditions support proportionate arrangements rather than escalation in demands.

Finally, the thesis focuses primarily on evidencing and reporting dynamics and only indirectly addresses deeper political questions about who gets to define impact and whose knowledge counts as credible. Future research could examine these power dimensions more explicitly. This could include how categories embedded in call templates shape which projects appear fundable, how intermediary support redistributes advantages and disadvantages, and how the politics of "proof" affects which forms of community value become visible.

Taken together, these directions suggest that the patterns identified here are best treated as a starting point. The thesis offers a grounded account of how evidence practices and evidence demands meet in small community-based settings, and it opens up questions for further work on how more proportionate, and more equitable, measurement arrangements could be developed and sustained.

8.4 Recommendations and Final Reflections

This thesis started from a tension that many projects in Brazil's third-sector ecosystem recognize in practice. "Impact measurement" has become a near-default language for legitimacy and funding, yet the everyday realities of small, community-based initiatives do not match the assumptions built into many measurement templates and reporting

routines. By placing project leaders' evidence practices alongside formal requirements embedded in CFPs, and by including specialist perspectives that move between these arenas, the study treats this tension not as an abstract debate about methods, but as a lived problem of translation, credibility, and fairness.

A first reflection is that the starting point for measurement debates matters. The empirical material challenges the idea that small projects are "empty" of evidence and must be taught evaluation from scratch. Instead, these initiatives already generate information that is closely linked to their purposes and are designed for proximity to the territory. It works because it circulates through relationships, between coordinators and participants, between projects and schools, and between families and community actors, and because it is interpreted in context. What can look "informal" from a distance can be highly functional for decision-making at ground level, precisely because it is embedded in delivery.

A second reflection is that demands for evidence are not simply technical preferences. They organize access and recognition. In the CFPs arena, evidence is designed for portability. It must travel across organizational boundaries and be legible to reviewers who do not know the territory, the participants, or the day-to-day constraints of delivery. This pushes requirements toward predefined indicators, structured monitoring plans, and document bundles that support comparability, traceability, and risk management. These requirements therefore have a governance function beyond learning. They help sponsors justify selections and manage accountability at a distance. As discussed above, one consequence is that "measurement" becomes not only something that happens after a grant is awarded, but also a filter that shapes who applies, who is considered fundable, and whose work becomes visible in formal portfolios. When requirements implicitly assume separated administrative roles, stable systems, and staff time for reporting cycles, misfit becomes predictable. Projects either divert scarce effort away from delivery, comply selectively, rely on external help, or self-exclude from opportunities altogether.

This leads to a broader reflection about where the burden of alignment is placed. A common response to misfit is to frame it as a capacity deficit on the side of small initiatives. The thesis suggests a different reading. Many tensions are structural. They arise because the purposes evidence serves in community practice are not identical to the purposes evidence serves in funding and accountability systems, and because legibility

pressures can escalate independently of feasibility. Treating the gap as a deficit story risks producing the outcome that many practitioners fear, evaluation as a parallel routine that absorbs time, looks professional on paper, yet remains weakly connected to the work it claims to describe. In this sense, over-demanding measurement can inadvertently encourage fragile claims. Organizations feel pushed to promise what they cannot robustly defend, or to present simplified signals as if they were comprehensive accounts of change.

A central contribution of the thesis has been to propose a practical way of describing these dynamics without collapsing into a binary of "rigorous versus not rigorous." The purpose–feasibility–legibility triangle reframes the question. What is the evidence meant to do. What can be sustained as a routine in this setting. Whom must the evidence be legible to, and in what form. This lens does not resolve deeper political questions about who defines impact, but it helps make trade-offs explicit. It also highlights that legibility is unevenly distributed. Two projects with similar practices can appear very different to external audiences depending on their access to administrative support, writing capacity, and intermediary assistance. Once this is acknowledged, proportionality becomes more than a technical recommendation. It becomes a fairness criterion, because it affects which organizations are rewarded and which are systematically disadvantaged.

From this perspective, the most promising path forward is not to advocate one flagship method, but to strengthen "conversion" practices that begin from what already exists and make it more portable without displacing it. The thesis's conversion approach can be read as a set of modest design principles. Stabilize a small evidence routine. Keep claims cautious and contribution-oriented. Combine different types of proof so that no single indicator has to carry the entire story. A concise theory of change can function as a practical anchor, helping projects prioritize a small set of outcomes that are both meaningful locally and observable with available tools. A short indicator summary can provide a legible snapshot without pretending to be a complete account. Structured narratives or trajectories can preserve context and mechanism, showing how changes unfold for different participants rather than only reporting aggregate counts. Partner confirmations and basic administrative "starter packs" can reduce friction where CFPs require proof of organizational capacity. Importantly, these adaptations are not about shifting projects' primary accountability away from participants and communities. They are about making existing evidence easier to recognize in institutional systems.

A further reflection concerns the role of intermediaries. The material suggests that translation is not merely an optional service. It operates as measurement infrastructure. Intermediaries reduce the cost of legibility by helping projects interpret CFPs, assemble documentation, align local descriptions with institutional language, and stabilize routines that can be sustained over time. This has implications for equity. If translation capacity is what allows evidence to travel, uneven access to it reproduces uneven access to funding, regardless of the underlying quality of community work. Building proportionate measurement arrangements therefore requires not only asking projects to improve documentation, but also investing in shared support structures that reduce barriers, especially during application windows and reporting cycles.

Finally, the thesis invites a more careful stance on what "better measurement" should mean in these settings. A field-wide push for standardization can be understandable for portfolio management, but if it is pursued without attention to feasibility and local mechanisms of change, it risks narrowing what counts as value. Many outcomes that matter in community projects, such as stability in routines, sustained engagement, trust, belonging, and strengthened relationships with adults and institutions, do not disappear because they are hard to quantify. They disappear from what gets recognized in formal reporting. A proportionality stance insists that measurement should not reward what is easiest to count at the expense of what is substantively significant for participants. In practice, this implies designing CFPs and reporting expectations so that they assume low administrative capacity when they claim to target grassroots initiatives, and allowing plural evidence formats that combine a small indicator set with structured qualitative accounts.

Taken together, these reflections support a closing argument. Impact measurement is most likely to strengthen community-based work when it starts from practice, stays proportionate to real constraints, and treats legibility as something to be negotiated rather than imposed. The most constructive role for measurement in this ecosystem is not to force small initiatives into the mold of large-program evaluation, but to enable credible accountability and learning that does not compete with the relationships and day-to-day delivery through which change is produced. In this sense, the question "can small projects measure impact?" is better reframed as a design question. Under what evidence arrangements do projects, funders, and intermediaries make impact credible enough to be accountable, while keeping it flexible enough to reflect how change happens in place.

APPENDIX

Appendix A. Summary of selected calls for proposals (CFPs) and application/reporting templates

This appendix provides an aggregated synopsis of requirements extracted from ten CFP document packages and associated annexes/templates included in the documentary analysis. To support anonymity and reduce traceability to specific programs and application contexts, the synopsis is presented in combined form and does not reproduce full document titles, sponsor names, or direct links. The synopsis was produced with AI assistance (large language model summarization) and curated to retain only elements directly relevant to the analysis.

The synopsis consolidates recurring requirements across three stages: (i) proposal-stage submissions (including documentary proof and project design elements), (ii) assessment and scoring criteria (including expectations around indicators and evaluation readiness), and (iii) implementation-stage monitoring and reporting duties (including cadence, verification expectations, and template/platform requirements). It is included as a transparency aid for the RQ2 documentary findings and the RQ3 alignment/misalignment discussion using the purpose–feasibility–legibility lens.

A.1 Documentary corpus overview (anonymized)

Table 2 summarizes the ten document packages in anonymized form. Labels (A–J) are used only for internal organization of the synopsis.

Document label (anonymized)	Document type / focus	Evidence / measurement expectations	Administrative / compliance expectations
CFP A	Public-sector call: implementation and reporting rules	Periodic activity/result reporting; combination of quantitative and qualitative evidence; final results narrative.	Formal annexes/declarations; financial reporting requirements; submission via a designated portal/system.
CFP B	Public utility / socioenvironmental call: selection and proposal content	Proposal scored on targets and indicators; requirement to state how targets will be measured; emphasis on coherence.	Strict checklist of organizational documents; disqualification rules for missing or incorrect submission.
CFP C	Public-sector call (sports/youth):	Actions, targets and indicators required; indicators that verify	Eligibility through registration/portal procedures;

	proposal minimum content and scoring	targets' treated as a scored item.	legal/fiscal documentation.
CFP D	Corporate call (sports): selection process + monitoring template	Explicit criterion on monitoring/evaluation approach; use of provided monitoring spreadsheet aligned to targets.	Multi-stage review; documentation deadlines; standard compliance package for selected organizations.
CFP E	Corporate/incentive program regulation: technical screening criteria	Adequacy of indicators and evaluation strategy assessed; some mandatory indicator categories used for comparability.	Team/compliance requirements (e.g., accounting); additional authorizations/licenses requested at later stages.
CFP F	Transport-sector corporate call: application guide prompts	Requires definition of indicators, verification sources, periodicity, and evaluation horizon.	Emphasis on internal consistency between methods, indicators, and budget lines; baseline organizational information.
CFP G	Small-grants application form (low-burden template)	Requests outcomes and indicators; suggests lightweight evidence (photos, lists, short feedback tools, testimonials).	Basic administrative information; data-protection commitments; partner rules when applicable.
CFP H	Public-sector call: mandatory documentation annex	Not a measurement tool itself; specifies proof items that enable later monitoring.	Extensive document list (statutes, registrations, certificates, financial statements, capacity attestations).
CFP I	Public-sector call: selection criteria annex	Scoring categories include coherence, feasibility, and monitoring/evaluation logic.	Scores organizational track record and transparency signals (e.g., prior reports, public communication, audits).
CFP J	Monitoring/reporting model template (standalone)	Predefined fields for outputs/outcomes, indicator values, evidence attachments, and progress notes.	Standardized reporting cadence and formatting; documentation of expenses and proof of delivery.

Table 2. Summary of Calls for Proposals used for analysis (own table).

A.2 Extraction categories used in the documentary analysis

Each document package was read in full and extracted into a common template, then coded along six categories that structure the results chapter:

- (a) Eligibility and legal/administrative requirements (who can apply; disqualification rules).
- (b) Project description, problem framing, and justification (diagnosis of the territory/problem).
- (c) Workplan, methods, targets and indicators (what will be done; how success is defined).
- (d) Budgets and financial control (eligible costs; accounting expectations).
- (e) Monitoring, reporting, verification, and platforms/systems (cadence, evidence attachments, submission procedures).
- (f) Selection criteria and scoring (how 'quality' and 'credibility' are operationalized in review).

A.3 What proposals are expected to contain (proposal stage)

Across the CFP corpus, proposal templates tend to standardize a similar minimum package. Applicants are expected to translate a context-specific problem into a structured intervention logic that is legible to external reviewers.

- Problem framing and territorial diagnosis: a short description of the local issue, target group, and why the intervention is needed.
- Objectives, activities and deliverables: a workplan that specifies actions, timeline, and responsibilities.
- Targets and indicators: explicit quantitative or qualitative targets and how indicators will be calculated.
- Verification approach: what evidence will be produced (e.g., lists, photos, logs, short instruments), who will produce it, and how often.
- Budget coherence: cost items aligned with activities and, where relevant, with monitoring/reporting tasks.

A.4 Eligibility and documentation burdens

A consistent cross-document feature is that eligibility is treated as a front-loaded screening step, often requiring a substantial bundle of organizational documentation before substantive project quality is considered.

- Proof of legal status and governance (e.g., statutes, board/management records, registrations).

- Fiscal and labor compliance certificates and declarations (often multiple, time-bounded documents).
- Financial statements and accounting documentation (e.g., balance sheet/income statement; accountant sign-off).
- Technical capacity attestations and track record evidence (prior projects, partner confirmations, past reporting).
- Portal/system registration requirements and standard forms that must be completed within strict deadlines.

For small community initiatives, these requirements can create a threshold effect: organizations may have meaningful delivery capacity but struggle to maintain the administrative infrastructure that makes capacity legible to reviewers.

A.5 How selection criteria operationalize ‘good evidence’

Selection tables typically translate broad expectations into scored criteria. Three recurring patterns are visible across the corpus:

A.5.1 Coherence and internal logic

Review criteria frequently reward proposals that present a clear chain from diagnosis to objectives, activities, targets and indicators. In practice, this privileges applications that can express local knowledge in a structured narrative format.

A.5.2 Indicators and evaluation as a scored item

Multiple documents treat the quality of indicators, monitoring plans, and ‘evaluation strategy’ as a distinct scored component. This operationalizes credibility as the ability to specify measurable targets and to define how evidence will be produced and verified.

A.5.3 Capacity and transparency signals

Scoring frameworks also reward organizational age, prior experience, and documentary signals of transparency (such as activity reports, public communication of results, and, where available, external audits). These criteria increase legibility for distant reviewers but can implicitly disadvantage initiatives with limited administrative support.

A.6 Monitoring and reporting duties during implementation

Implementation requirements vary in intensity, but the corpus shows a shared expectation that funded organizations will produce periodic reports combining indicator values with evidence attachments. Report cadence ranges from regular (e.g., monthly or quarterly) to milestone-based or end-of-project reporting, depending on program design.

Common evidence formats requested include:

- Attendance and participation records (lists, enrolment forms, session logs).
- Photos/audiovisual documentation of activities and outputs.
- Spreadsheets aggregating outputs, beneficiaries reached, and progress toward targets.
- Short feedback instruments (e.g., satisfaction questionnaires) and, where feasible, simple pre/post comparisons.
- Narrative evidence (testimonials, short case notes, or trajectory descriptions) to document qualitative change.

Several documents also frame reporting as a compliance routine, linking evidence submission to financial accountability and to continued eligibility for payments or follow-on support.

A.7 Relevance to the thesis analytical lens

Taken together, the documentary corpus illustrates how templates and criteria translate ‘impact’ into legible administrative and measurement routines. These routines often prioritize portability (comparability, auditability, standardized reporting) and therefore create pressures on feasibility for low-resource initiatives. The synthesis supports the thesis’ discussion of purpose (why evidence is demanded), feasibility (what can realistically be sustained), and legibility (what formats travel in funding systems).

Appendix B. **Interview Guide - Project Leaders**

Interview guide – question script

Note: avoid proper names and identifying information. Whenever possible, ask for concrete examples. Questions about funding can be skipped if the interviewee does not feel comfortable.

Core section – questions for all interviewees

1. Project and role

- What is the name of the project? (A short nickname is fine.)
- What is your role in the project?
- What is the project’s main objective?

2. Start and how it works

- When did the project start?
- How does the project work in practice? (main activities, how people join, etc.)

3. Team (approximate)

- Today, about how many people help run the project?
- Roughly, how many are paid staff and how many are volunteers?

4. People served (no names)

- Who usually takes part in the activities?
- Is there any participation criteria? (if applicable)

5. Participation frequency and duration

- When someone joins the activities, how often do they participate and how long do they usually stay involved?
- If helpful, give an example (e.g., once a week for 3 months; twice a week during the school term).

6. Partnerships (types only)

- What types of partners do you work with? (e.g., schools, health posts/clinics, churches, neighborhood groups, NGOs, companies).
- No need to mention names.

7. What does “good results” mean here?

- When you say “it’s working”, what changes do you see in people or in the community?
- What goal(s) need to be achieved for the project to be “working”?
- Who tends to benefit first?
- Could you share two short stories from the last year that made you think: “Yes, the project helps.”
- Without using names; just tell what happened (a few lines are enough).

9. Funding and reporting (optional)

- If you feel comfortable answering: have you received funding or any kind of support for the project? (Naming supporters is optional.)
- Before providing support, what do supporters typically ask for? (e.g., a form, certain numbers, photos, a plan).
- After support is provided, what do they usually ask for? (e.g., monthly/quarterly report, numbers, short stories, photos, receipts).
- If you have more than one supporter, do they ask for different things?
- Do supporters usually follow up on the project’s development? Do they ask informally or require documents, photos, reports, etc.?
- How did you connect with this/these supporter(s)?

10. Main difficulty in demonstrating results (choose one option and give an example)

- ☐ Purpose: we are not sure what is useful to communicate (e.g., “We don’t know which 2–3 things would help us tell our story.”)
- ☐ Feasibility: not enough time/people/tools (e.g., “We don’t have hours, staff, a phone, internet, or a simple sheet to record activities.”)

- ☐ Legibility: it is hard to match what supporters ask for (e.g., “They ask for forms or numbers we don’t use day to day, or each supporter asks for something different.”)

11. Current records (to direct the track)

- Do you currently record or keep anything about activities or results? Even something simple, such as attendance lists, short notes, or photos (with permission).
- If yes, go to Track M.
- If no, go to Track N.

Track M – for projects that already record something

M1. What do you record today?

- Tell me what you actually record or keep in practice (e.g., attendance, age range, simple “before and after” questions, short stories, photos).
- Only include what you really keep.

M2. Why do you record it?

- Why do you record it? Is it mainly to learn within the team, to report to supporters, or to decide next steps?
- It can be more than one reason.

M3. Where do you store it and how often do you review it?

- Where do you keep this information (e.g., paper, WhatsApp, Excel/Google Sheets, an app)?
- How often do you look back/review it (e.g., after each activity, monthly, only when there is a report due, or almost never)?

M4. Small and early signs

- What are the small, early signs that the project is helping?
- Examples: people keep coming back, someone returns after a break, someone accepts a referral, a teacher mentions improved attendance.

M5. Fit with what supporters ask for (0–10)

- On a scale from 0 to 10, how well do current requests fit with what you can collect honestly? (0 = not at all; 10 = fits well).
- What change would make that fit better?

Track N – for projects that do not yet record systematically

N1. Is there any simple record (even informal)?

- Do you keep any simple record, even if it is messy? (e.g., a list of names, attendance lists, notes on paper, WhatsApp notes).

N2. Main reasons for not recording/measuring yet

- What are the main reasons you do not yet have a way to record something or show any results?
- Examples: lack of time, small team, not knowing where to start, privacy concerns, difficulty getting information from partners, nobody asks.

N3. Small and early signs (even without formal records)

- Even without formal notes, what signs show that the project helps?
- Examples: people keep coming, bring a friend, accept a referral, fewer conflicts, someone returns to school.

N4. If you were to record only 3 things for 6 months

- If you had to choose only three simple things to record over the next 6 months, what would they be?
- If you want, you can use examples (e.g., how many come back every week; how many return after a break).

N5. What would help you start recording

- What kind of support would help you start recording something?
- Examples: a very simple sheet, a one-page plan on "how change happens", a very short feedback form, a short training, a volunteer to organize notes.

N6. A "light" option if someone asks for something heavy

- If someone asked for a heavy report or a single "money number", what lighter alternative would work?
- E.g., 3 simple counts + 1 very short story per month.

N7. Information from partners (simple and realistic)

- Is there any simple number a partner could share (with permission) that would help you? (e.g., the school confirming that a student returned).
- What gets in the way today: format, permission rules, lack of time, something else?

Closing – for all interviewees

F1. One thing you really wish you could show

- What is one thing you do not record today, but that, if you could show it, would make people truly understand the project's value?

F2. Anything missing?

- Is there anything important I did not ask that you think I should know?

Appendix C. Interview guide – Specialists

1. Warm-up and perspective

- Could you briefly describe your role and your experience with small, community-based social projects?

- Probe: In what type of work do you most often engage with these projects, impact evaluation, consulting/advisory, institutional strengthening (capacity building), support for project design/fundraising, research, or something else?

2. Complex problems and impact expectations

- When you think about the kinds of social problems these small projects work on, how realistic is it to expect a clear, linear cause-and-effect relationship between a single project and the final impact?
- In your practice, how do you distinguish between a reasonable expectation of "demonstrating impact" for a small project and a requirement that is already unrealistic?

3. Evidence in practice and credibility

- When small, resource-constrained projects present their results to you, what do you pay most attention to in deciding whether the evidence is "good enough" to be considered credible?

- Probe: Are there specific signals or patterns (e.g., consistency over time, triangulation of sources, partner voices, plausibility of the narrative) that matter more than the format itself?

- Many grassroots projects rely on things like testimonials, photos, partner feedback, attendance lists, and simple counts. In your view, what are the main strengths and limitations of this kind of "everyday evidence"?
- When you need to assess more intangible outcomes (e.g., self-esteem, social cohesion, discipline, future aspirations), how do you like to combine quantitative elements (records, short scales, counts) with qualitative elements (stories, observations, partner voices)?

- Optional: Can you recall a case where combining simple numbers with well-documented stories gave you a much better understanding of a project's effects?

4. Tools and frameworks in small projects

- In your work with small organizations or community-based groups, how often do tools like log frames, theories of change, SROI, IRIS+, or similar frameworks come up, and in what kinds of situations does this typically happen?

- Based on your experience, what tends to work well when these frameworks are used by very small teams, and what tends to be unrealistic or even counterproductive?

- Probe: Are there specific steps (e.g., mapping outcomes, defining indicators, monetization, counterfactual) that you consider especially problematic in low-resource contexts?

- If you had to adapt these more formal tools to something proportional to the reality of a very small, volunteer-led project, what would you keep and what would you leave out?

5. Proportionate evaluation and the ecosystem

- If you had to design a minimal evaluation structure for a small project, one that still respects the complexity of the context it operates in, what elements would you definitely include?

- Probe: For example, a simple results chain, 1–2 core outcome indicators, a basic participant register, periodic reflection moments about the context, partner feedback, etc.

- Thinking about how social projects are seen and funded in Brazil, what types of organizations tend to become more "visible" and attractive to funders, and what types of grassroots initiatives tend to remain invisible, even when they are doing good work? What, in your view, creates this difference?
- What kinds of support or arrangements (e.g., networks, shared tools, intermediary organizations) help small projects evaluate themselves in a way that is both realistic and consistent with the complexity of the problems they face?

Glossary

Term	Definition
Accountability	Expectation that projects justify the use of resources and explain what was delivered or achieved to funders, partners, or the public. Often operationalized through reporting requirements and documentation.
Analytic note (memo)	Short written note produced during analysis to capture interpretations, emerging patterns, and links to the research questions.
Audit trail	Record of key analytic and procedural decisions, such as what was included or excluded and how interpretations were reached. It allows readers to follow how conclusions were produced.
Call for Proposals (CFP) / “Editais”	Funding call that sets eligibility rules, required documents, and proposal and reporting formats that applicants must follow to receive and retain funding.
Capability approach	Perspective that assesses change through people’s freedoms to do and be what they value, not only through services delivered.
Coercive pressures	Requirements from authority, such as legal rules or reporting clauses, that push organizations to adopt particular practices to secure funding or legitimacy.
Community-based social project	Local initiative delivering social support or services with small teams and limited administrative capacity, typically working close to residents and local partners.
Counterfactual	Estimate of what would have happened without the intervention. Used as the comparison scenario in causal impact evaluation.
Difference-in-differences	Quasi-experimental method that estimates an intervention effect by comparing change over time in a treated group to change over time in a comparison group.
Evidence (impact evidence)	Any information used to support claims about activities, outputs, outcomes, or change. In this thesis it includes routine artefacts and funder-requested indicator reporting.

Feasibility	Whether an evidence practice can be sustained in low-resource routines given time, staff skills, data access, funding stability, and administrative workload.
Field (organizational field)	Network of actors, such as funders, NGOs, intermediaries, consultants, and regulators, whose interactions create shared expectations about proper practices.
Fit (purpose–feasibility–legibility fit)	Degree of alignment between what evidence is for (purpose), what can realistically be produced (feasibility), and what is recognized as credible by relevant audiences (legibility).
Indicator	Defined measure used to track progress or results, such as participants reached, completion rates, or change in a specific outcome.
Institutional complexity	Situation where organizations face competing logics or expectations, for example learning versus accountability, and must navigate trade-offs or hybrid responses.
Institutional pressures	Forces that shape what counts as valid evidence and practice in a field, including coercive, normative, and mimetic pressures.
Intermediaries	Actors that help projects translate local documentation into funder-facing formats. They often provide guidance, templates, and capacity support at a cost.
Legibility	Whether evidence is recognizable, understandable, and seen as credible by external audiences, especially funders, in selection and reporting settings.
Legibility labor	Extra work needed to make locally produced information travel into external reporting systems. This can include formatting, translating, aggregating, and aligning with templates.
Loose coupling	Gap between formal reporting structures and day-to-day practice. Compliance systems exist but are only weakly connected to delivery routines.
“Marco lógico” (logical framework)	Planning and evaluation matrix that links objectives, activities, outputs and outcomes, indicators, verification sources, assumptions, and risks.

Mimetic pressures	Pressures to copy widely recognized practices under uncertainty, such as adopting standard indicator sets because they are common.
Monitoring and evaluation (M&E)	Activities to track implementation and results over time. It ranges from routine monitoring to evaluation of outcomes for learning and accountability.
Normative pressures	Pressures arising from professional norms and guidance, such as evaluation standards and best-practice toolkits, that define what good measurement looks like.
Outcome	Change experienced by participants or systems that is plausibly linked to project activities, such as skills gained, increased access, or behavior change.
Portability	Ability of evidence to travel across contexts and audiences in a standardized form, such as dashboards or indicator tables, that is easy to interpret externally.
Proportionality	Principle that measurement demands should be scaled to the decisions at stake and the capacity available, avoiding burdens that exceed what is useful or feasible.
Propensity score	Estimated probability of receiving an intervention given observed characteristics. Used to create more comparable groups in observational studies, for example through matching or weighting.
SROI (Social Return on Investment)	Approach that monetize outcomes to estimate social value created per unit of investment, comparing benefits to costs.
Theory of Change	Explanation of how and why an intervention is expected to lead to outcomes, including assumptions and causal pathways.
Third sector	Organizations outside government and for-profit business, such as NGOs, institutes, associations, and foundations, that pursue public or social purposes.
Validity and reliability (indicators)	Validity means an indicator measures what it claims to measure. Reliability means it measures consistently across time, observers, or contexts.

Bibliography

Banerjee, S. B. (2008). Corporate social responsibility: The good, the bad and the ugly. Critical Sociology, 34(1), 51–79.

Barnett, M. L., Henriques, I., & Husted, B. W. (2020). Beyond good intentions: Designing CSR initiatives for greater social impact. Journal of Management, 46(6), 937-964.

Baxter, P., & Jack, S. (2008). Qualitative case study methodology: Study design and implementation for novice researchers. The Qualitative Report, 13(4), 544–559.

Bertrand, M., Duflo, E., & Mullainathan, S. (2002). How much should we trust differences-in-differences estimates? (NBER Working Paper No. 8841). National Bureau of Economic Research. https://doi.org/10.3386/w8841

Brent, R. J. (2023). Cost-Benefit Analysis versus Cost-Effectiveness Analysis from a Societal Perspective in Healthcare. International Journal of Environmental Research and Public Health, 20(5), 4637. https://doi.org/10.3390/ijerph20054637

Conselho Nacional de Saúde. (2016). Resolução nº 510, de 07 de abril de 2016. Ministério da Saúde. https://www.gov.br/conselho-nacional-de-saude/pt-br/atos-normativos/resolucoes/2016/resolucao-no-510

Crowe, S., Cresswell, K., Robertson, A., Huby, G., Avery, A., & Sheikh, A. (2011). The case study approach. BMC Medical Research Methodology, 11, 100. https://doi.org/10.1186/1471-2288-11-100

Câmara dos Deputados. (2018). Lei nº 13.709/2018 , Lei Geral de Proteção de Dados Pessoais (LGPD). https://www2.camara.leg.br/legin/fed/lei/2018/lei-13709-14-agosto-2018-787077-normaatualizada-pl.pdf

Davies, R., & Dart, J. (2005). The ‘Most Significant Change’ (MSC) technique: A guide to its use (Version 1.00). University of Auckland. https://cdn.auckland.ac.nz/assets/auckland/education/research/docs/CCRE-MSCGuide.pdf

DiMaggio, P. J., & Powell, W. W. (1983). The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields. American Sociological Review, 48(2), 147–160.

Donati, P. (2014). Relational goods and their subjects: The ferment of a new civil society and civil democracy. Recerca. Revista de Pensament i Anàlisi, 14, 19–46.

Ebrahim, A., & Rangan, V. K. (2014). What impact? A framework for measuring the scale and scope of social performance. California Management Review, 56(3), 118–141.

Feor, L., Clarke, A., & Dougherty, I. (2023). Social impact measurement: a systematic literature review and future research directions. World, 4(4), 816-837.

Gale, N. K., Heath, G., Cameron, E., Rashid, S., & Redwood, S. (2013). Using the framework method for the analysis of qualitative data in multi-disciplinary health research. BMC Medical Research Methodology, 13, 117.

Gertler, P. J., Martinez, S., Premand, P., Rawlings, L. B., & Vermeersch, C. M. J. (2016). Impact evaluation in practice (2nd ed.). World Bank & IDB.

Global Impact Investing Network (GIIN). (2019). IRIS+ Core Metrics Sets: Overview. GIIN.

Global Impact Investing Network (GIIN). (2022). Sizing the impact investing market.

Global Reporting Initiative. (2021). GRI 1: Foundation 2021. https://globalreporting.org/pdf.ashx?id=12334

Greenwood, R., Raynard, M., Kodeih, F., Micelotta, E. R., & Lounsbury, M. (2011). Institutional complexity and organizational responses. Academy of Management Annals, 5(1), 317–371.

GRI, UN Global Compact, & WBCSD. (2015). SDG Compass: The guide for business action on the SDGs.

Hassnain, H., Kelly, L., & Somma, S. (2021). Evaluation in contexts of fragility, conflict and violence: Guidance from global evaluation practitioners.

Hennink, M., & Kaiser, B. N. (2022). Sample sizes for saturation in qualitative research: A systematic review of empirical tests. Social Science & Medicine, 292, 114523.

Holmes, A. G. D. (2020). Researcher positionality, A consideration of its influence and place in qualitative research: A new researcher guide. Shanlax International Journal of Education, 8(4), 1–10.

Howaldt, J., & Schwarz, M. (2010). Social innovation: Concepts, research fields and international trends. Dortmund: IMO.

Impact Economy Foundation. (2024). Summary: Impact-Weighted Accounts Framework (IWAF). Impact Economy Foundation.

Insper Metricis. (2018). Guia de avaliação de impacto socioambiental para utilização em negócios e investimentos de impacto: Guia geral com foco em verificação de adicionalidade (4ª rev.). São Paulo: Insper.

Kah, S., & Akenroye, T. (2020). Evaluation of social impact measurement tools and techniques: a systematic review of the literature. Social Enterprise Journal, 16(4), 381-402.

Kwizela, R., Dugange, A., Kabole, I., Murungu, R., & Watako, D. (2018). Building Entrepreneurship for Water, Liquid and Solid Waste Management in Temeke

Municipal Council of Dar-es-Salaam Tanzania: A Social Return on Investment Analysis.

Lacerda, D. S. (2014). Sobre o impacto social e sua avaliação em organizações baseadas na favela. Favelas at LSE

Lazzarini, S. G., Setter Filho, J. G., Ikawa, J. N. R., & de Barros, O. A. D. (2021). Monetização de impacto social: Análise comparativa de ferramentas alternativas e sua aplicabilidade. São Paulo: Insper Metricis.

Lazzarini, S. G., Setter Filho, J. G., Melo, C. P. G. de, Ikawa, J. N. R., de Barros, O. A. D., & Castejon, C. P. (2022). Guia de avaliação de impacto socioambiental para utilização em projetos e investimentos de impacto: Guia geral com foco em monitoramento e verificação de adicionalidade (5ª ed.). São Paulo: Insper Metricis.

Maguire, M., & Delahunt, B. (2017). Doing a thematic analysis: A practical, step-by-step guide for learning and teaching scholars. AISHE-J: The All Ireland Journal of Teaching and Learning in Higher Education, 8(3), Article 3351.

Malterud, K., Siersma, V. D., & Guassora, A. D. (2016). Qualitative interview studies: Guided by information power. International Journal of Qualitative Studies on Health and Well-being, 11, 30996.

Meyer, J. W., & Rowan, B. (1977). Institutionalized organizations: Formal structure as myth and ceremony. American Journal of Sociology, 83(2), 340–363.

Millar, R., & Hall, K. (2013). Social Return on Investment (SROI) and performance measurement: The opportunities and barriers for social enterprises in health and social care. Public Management Review, 15(6), 923–941.

Mishra, B. (2018). Social Impact Measurement and Investment: Methods, Limitations and Challenges. Transcience, 9(1), 20-47.

Monteiro, H. (2010). Small grants in a big country. Alliance, 15(1), 53–54.

Mulgan, G. (2006). The process of social innovation. Innovations: Technology, Governance, Globalization, 1(2), 145–162.

New Economics Foundation. (2009). Measuring value: A guide to social return on investment (SROI) (2nd ed.).

Nicholls, J., Lawlor, E., Neitzert, E., & Goodspeed, T. (2012). A Guide to Social Return on Investment (The SROI Guide). The SROI Network.

Noble, H., & Smith, J. (2015). Issues of validity and reliability in qualitative research. Evidence-Based Nursing, 18(2), 34–35.

Nowell, L. S., Norris, J. M., White, D. E., & Moules, N. J. (2017). Thematic analysis: Striving to meet the trustworthiness criteria. International Journal of Qualitative Methods, 16, 1–13.

Nussbaum, M. C. (2011). Creating capabilities: The human development approach. Cambridge, MA: Harvard University Press.

O'Brien, B. C., Harris, I. B., Beckman, T. J., Reed, D. A., & Cook, D. A. (2014). Standards for reporting qualitative research (SRQR): A synthesis of recommendations. Academic Medicine, 89(9), 1245–1251.

Palinkas, L. A., Horwitz, S. M., Green, C. A., Wisdom, J. P., Duan, N., & Hoagwood, K. (2015). Purposeful sampling for qualitative data collection and analysis in mixed method implementation research. Administration and Policy in Mental Health and Mental Health Services Research, 42(5), 533–544.

Prates Rodrigues, M. C. (2010). Projetos sociais corporativos: como avaliar e tornar essa estratégia eficaz. São Paulo, SP: Atlas.

Prates Rodrigues, M. C. (2011). Organizações do terceiro setor e projetos sociais. GIFE – Grupo de Institutos, Fundações e Empresas. Retrieved July 3, 2025, from https://gife.org.br/organizacoes-do-terceiro-setor-e-projetos-sociais/

Prates Rodrigues, M. C. (2014). Planejamento e avaliação de projetos sociais: o marco lógico revisitado. GIFE – Grupo de Institutos, Fundações e Empresas. https://gife.org.br/planejamento-e-avaliacao-de-projetos-sociais-o-marco-logico-revisitado/

Prates Rodrigues, M. C. (2018). Avaliação de projetos sociais no terceiro setor: Uma agenda em construção. Revista Brasileira de Monitoramento e Avaliação, 10(1), 72–89.

Prates Rodrigues, M. C. (2021). Avaliação de impacto: fantasia ou realidade? Portal do Impacto. Retrieved July 2, 2025, from https://www.portaldoimpacto.com/avaliacao-de-impacto-fantasia-ou-realidade

Prates Rodrigues, M. C. (2024). Exigir medição de impacto e SROI para o terceiro setor? Estratégia Social. Retrieved July 2, 2025, from https://estrategiasocial.com.br/exigir-medicao-de-impacto-e-sroi-para-o-terceiro-setor/

Presidência da República. (2006). Lei nº 11.438, de 29 de dezembro de 2006. https://www.planalto.gov.br/ccivil_03/_ato2004-2006/2006/lei/l11438.htm

Presidência da República. (2013). Decreto nº 7.988, de 17 de abril de 2013 (PRONON e PRONAS/PCD). https://www.planalto.gov.br/ccivil_03/_ato2011-2014/2013/decreto/d7988.htm

Presidência da República. (2022). Lei nº 14.439, de 24 de agosto de 2022. https://www.planalto.gov.br/ccivil_03/_ato2019-2022/2022/lei/L14439.htm

Receita Federal do Brasil. (2024). Destinação de imposto de renda por empresas [Folder]. https://www.gov.br/receitafederal/pt-br/centrais-de-conteudo/publicacoes/folheteria/destinacao-de-imposto-de-renda-por-empresas-folder.pdf

Receita Federal do Brasil. (2025). Destinação como Pessoa Jurídica (resumo oficial de limites; requisito de lucro real; total possível de 11%). https://www.gov.br/receitafederal/pt-br/acesso-a-informacao/acoes-e-programas/cidadania-fiscal/destinacao-irpf/destinacao-como-pessoa-juridica

Receita Federal do Brasil. (n.d.). Projetos culturais (Lei Rouanet): Incentivo fiscal (pessoas físicas e jurídicas tributadas pelo lucro real). Retrieved October 31, 2025, from https://www.gov.br/receitafederal/pt-br/acesso-a-informacao/acoes-e-programas/cidadania-fiscal/destinacao-irpf/projetos-culturais-lei-rouanet

Reeder, N., & Colantonio, A. (2013). Measuring impact and non-financial returns in impact investing: A critical overview of concepts and practice (EIBURS Working Paper 2013/01). LSE Cities. https://eprints.lse.ac.uk/59126/

Rogers, P. (2014). Theory of Change (Methodological Briefs: Impact Evaluation No. 2). UNICEF Office of Research.

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. Biometrika, 70(1), 41–55.

Serafeim, G., Zochowski, R., & Downing, J. (2019). Impact-Weighted Accounts Initiative: Measuring the total impact of companies. Harvard Business School.

Silveira, T. G. (2020). Avaliação e monitoramento de projetos sociais. Serviço Social & Realidade, 29(1), 75–92.

So, I., & Staskevicius, A. (2015). Measuring the “impact” in impact investing. Harvard Business School Social Enterprise Initiative. https://www.hbs.edu/socialenterprise/Documents/MeasuringImpact.pdf

Stephanou, L., Müller, L. H., & Carvalho, I. C. de M. (2003). Guia para elaboração de projetos sociais (2nd ed.). São Leopoldo: Editora Sinodal; Fundação Luterana de Diaconia.

Tong, A., Sainsbury, P., & Craig, J. (2007). Consolidated criteria for reporting qualitative research (COREQ): A 32-item checklist for interviews and focus groups. International Journal for Quality in Health Care, 19(6), 349–357.

Vanclay, F. (2003). International principles for social impact assessment. Impact assessment and project appraisal, 21(1), 5-12.

Vanclay, F., Esteves, A. M., Aucamp, I., & Franks, D. (2015). Social Impact Assessment: Guidance for assessing and managing the social impacts of projects. International Association for Impact Assessment.

Vogel, I. (2012). Review of the use of ‘Theory of Change’ in international development. UK Department for International Development (DFID).

Wilson-Grau, R., & Britt, H. (2013). Outcome Harvesting (Brief). Ford Foundation MENA Office.