Original Article

Methodological Innovations
1–18

DOI: 10.1177/20597991251383745
journals.sagepub.com/home/mio
Sage

Social responsibility and its impact: A methodological proposal

Matilde Valente Rosa[1,2], Leonor Bacelar-Nicolau[2] and Maria de Fátima Ferreiro[3]

Abstract

Social responsibility has emerged as one of the predominant topics in public discourse since the beginning of the century, not only at the level of social organizations but also within corporate spheres. Since the 1990s, many methods have been developed to measure social impact, showing the extent to which there is a need to evaluate this complex and multidimensional reality. However, despite these multiple efforts, none of these methodologies has become a reference standard. An exhaustive literature review was conducted to carry out a classification of SIA models, aimed at understanding their strengths and weaknesses, as well as identifying any gaps within the sector. The work encompassed both mapping and classifying SIA models along with identifying indicators related to these models. Later, efforts were made to harmonize indicator nomenclature through a combination of manual and automated methods. 144 methodologies and 1361 indicators were listed. First, the study concluded that there is no universally accepted definition of the concept of Social Impact. Second, the study concluded that no existing measurement fully meets the three fundamental characteristics outlined in the literature: **producing a quantitative output**, **being exhaustive** (i.e., considering all stakeholders), and **enabling comparability** over time and across (social and business) organizations or projects with differing characteristics.

Keywords

Social responsibility, social impact, measurement models, classification, text mining, social impact indicators

Introduction

Since the beginning of the 21st century, the discourse surrounding social responsibility has emerged as a paramount focal point within the sphere of social organizations but also within the corporate sector (Corvo et al., 2021). Some European guidelines, namely those issued by the European Commission (2014a), have reinforced the importance of this topic in society, mandating that large entities start reporting on their environmental and Social Impact (SI) within their management reports. Among those European guidelines, two of them can be highlighted. The EU Sustainable Finance Disclosure Regulation (SFDR) 2019/2088 aims to integrate ESG disclosure and reduce information asymmetries in the financial services sector (Bengo et al., 2022). Regulation EU 2020/852 establishes a common language for sustainability and aligns criteria for determining sustainable economic activities. In particular, article 8 focuses on the disclosure of information about financial products that respect and meet socio-environmental characteristics, including the methodology used to measure these characteristics (Bengo et al., 2022).

This involves the acknowledgment of value beyond monetary value and corresponds to a paradigmatic change towards a broad approach of an enterprise's activity. The proliferation of organizations dedicated to addressing complex

[1]Department of Information Sciences and Technologies University: ISCTE-IUL, Instituto Universitário Lisboa, Lisboa, Portugal
[2]ISAMB-Institute of Environmental Health, Faculdade de Medicina, Universidade de Lisboa, Lisboa, Portugal
[3]Department of Political Economy University: ISCTE-IUL, Instituto Universitário Lisboa, Lisboa, Portugal

Leonor Bacelar-Nicolau is also affiliated to Universidade Católica Portuguesa, Católica Medical School, Center for Interdisciplinary Research in Health, Portugal.

Corresponding author:
Matilde Valente Rosa, Department of Information Sciences and Technologies University: ISCTE-IUL, Instituto Universitário Lisboa, Av. das Forças Armadas, Lisboa 1649-026, Portugal.
Email: matilde_rosa@iscte-iul.pt

social issues that have been historically neglected or inadequately addressed by public entities explains the emphasis on SI and the growing interest on social entrepreneurship and its impact (Kah and Akenroye, 2020). This is also the case of the academic-scientific community that presented different studies on the subject (Bernardino and Santos, 2020).

This paper aims to conduct a benchmark of Social Impact Assessment (SIA) models, highlighting their characteristics, strengths, and weaknesses. The objective is twofold: first, to assist professionals in the field in selecting the most suitable model based on their specific needs. Second, to determine if a single method encompasses the most important characteristics outlined in literature.

There are many SIA models, with different assumptions, leading to results that become incomparable between organizations or between types of beneficiaries. Drawing upon the literature review, three core characteristics have been identified as essential to a robust approach to social impact assessment: the quantitative dimension, comparability, and exhaustiveness. These characteristics reflect fundamental requirements consistently emphasized in the field for conducting effective and credible evaluations. Collectively, these characteristics serve as foundational pillars for comprehensive and methodologically sound impact measurement.

(1) **Quantitative dimension** – obtaining, with each measurement, a final impact value, numerical and contained in a range of possible values. This global final value can be disaggregated into sub-dimensions and must have an intelligible interpretation (Grieco et al., 2015);
(2) **Comparability** – allowing benchmarking between different types of projects, different types of stakeholders, different types of organizations, and different periods in time (Molecke and Pinkse, 2017);
(3) **Exhaustiveness** – measuring all aspects of a social initiative – whether undertaken by social organizations or companies – referring to the object (direct and indirect beneficiaries, employees, partners, environment, etc.) and the area of intervention (well-being, skills, personal fulfillment, etc.) (Costa and Pesci, 2016).

The paper is organized into the following sections: the second section presents a review of the literature on the subject, including methodologies for SIA, their main characteristics, and limitations; section three outlines the methodological options employed in the research; section four presents the results and discusses the findings. The final section considers the results in the context of the existing literature, culminating in the conclusion of the study.

Review of the literature

Humanity faces increasingly complex and interrelated challenges concerning production and consumption systems, including climate change, biodiversity loss, social inequality, and the overexploitation of natural resources. In this context, sustainable entrepreneurship emerges not only as a driver of value creation but also as a catalyst for systemic change. It plays a critical role in promoting environmental resilience, advancing inclusive and equitable economic development, and fostering innovation that contributes to the achievement of long-term sustainability goals (Matzembacher et al., 2019).

In the second half of the 20th century, managers began to assume the responsibility of balancing the maximization of their profits with the creation of social value in the communities in which they operate (Carroll, 2009). Managers have begun to recognize that high Environmental, Social, and Governance (ESG) performance is associated with better financial returns (Bengo et al., 2022). The discourse surrounding companies' intentions regarding Corporate Social Responsibility (CSR) began, delineating perspectives spanning from profit-centric orientations to those prioritizing the generation of social value (Maas and Liket, 2011).

Companies observe a transformation in the discourse surrounding SI and evaluation methodologies, signifying an evolution in both expectations and the linguistic frameworks employed (Arvidson and Lyon, 2014). Enterprises promptly recognized that the concept of value extends beyond conventional financial goals (Corvo et al., 2021). Therefore, a shift has been observed from the Single Bottom Line (economic value) paradigm to a Multiple Bottom Line approach that encompasses considerations across economic, social, and environmental dimensions (Maas and Liket, 2011).

The Multiple Helices framework also helps to understand the multidimensional nature of social impact. Originating from the Triple Helix model, which emphasizes collaboration between academia, industry, and government, later expansions such as the Quadruple and Quintuple Helix add the roles of civil society and the environment, respectively. These models underscore that sustainable social impact is not generated in isolation but emerges through dynamic interaction among key societal sectors (Peris-Ortiz et al., 2016).

From the perspective of social organizations and companies engaged in socially responsible or green innovation projects, the exigency of complex social and environmental challenges – combined with competition for financial and human resources – drives the need for greater efficiency, effectiveness, and impact orientation. While social organizations often lead in addressing unmet societal needs, an increasing number of private firms are integrating social and environmental objectives into their business models, aligning with ESG criteria and sustainable innovation frameworks (Bernardino and Santos, 2020).

In summary, there is a growing recognition that attentiveness to social issues can yield favorable effects on corporate financial performance and a growing number of social investors and mutual funds use social criteria to evaluate their investments (Epstein and Roy, 2001).

The emphasis on SI has experienced an escalation, driven by the promulgation of policy documents, legislative initiatives, and an intensified focus on achieving the well-known sustainable development goals (Kah and Akenroye, 2020). The academic-scientific community has also demonstrated a growing interest in social entrepreneurship and its impact on society, resulting in a significant increase in publications that explore the SI of social entrepreneurship initiatives (Bernardino and Santos, 2020).

SIAs (SIA) have emerged as a discrete discipline within the broader field of impact assessment, concentrating on the social dimensions of sustainable development and evaluating the ramifications of planned interventions on communities (Arce-Gomez et al., 2015). Measuring SI must evolve into an integral component of an organization's operations and decision-making processes (European Commission, 2015).

SIA spans a spectrum ranging from qualitative, subjectively derived measures to multidimensional quantitative metrics (Öncer, 2019). Measurement must be relevant, useful, simple, natural, certain, understood and accepted, transparent, well-explained and supported by evidence. SI should cover the impact on communities, both short and long-term, and distinguish between direct and indirect SI (European Commission, 2014b). SIA must offer a thorough evaluation of SI, encompassing diverse dimensions such as demographics, individual well-being, community welfare, employment dynamics, and job satisfaction. Furthermore, SIA needs to be exhaustive in addressing the various stakeholders involved, encompassing evaluations of the impact on employees, suppliers, customers, and the broader local community (Miller et al., 2007). SIA aims to consider qualitative and quantitative aspects, addresses proportionality and relevance, and adapts to the evolving nature of social enterprises (European Commission, 2014b).

Evaluation serves as a valuable tool for leaders to demonstrate the positive impact of interventions (Lee et al., 2021). Demonstrating the value created and being accountable for actions and performance are crucial aspects for social initiatives (Bernardino and Santos, 2020). Assessments serve as a tool to identify, measure, and demonstrate the net benefits or damages resulting from an organization's activities, as well as the effects on its stakeholders (Barby et al., 2021).

Recent organizations tend to place greater value on evaluating results, realizing benefits such as greater effectiveness and the efficient use of resources (Bernardino and Santos, 2020), visibility, the increase of the likelihood of attracting financial support for the project (Lee et al., 2021), and future funders and donations (Arvidson and Lyon, 2014) and attraction of employees and volunteers (LeRoux and Wright, 2010). The positive effects on employee motivation, retention and recruitment, cost savings and increased revenue, customer attraction and retention, reputation, and employer attractiveness are clear (Weber, 2008). Among the main motivations for measuring impact are the improvement in demonstrating transparency, responsibility, and organizational legitimacy to stakeholders (including investors) (Lall, 2019).

Firms undertaking green innovation initiatives benefit from SIA models by systematically evaluating the environmental and social outcomes of their projects. These models support the identification of long-term value creation, enhance sustainability reporting, and provide measurable evidence of impact to regulators, investors, and customers. By integrating SIA into their innovation processes, companies can align their strategies with global sustainability goals and strengthen their competitive advantage in environmentally conscious markets (Liu et al., 2024).

The field of SIA has evolved over time, with contributions from the academia, international organizations, and professionals, leading to the development of multiple methodologies for impact assessment (Esteves et al., 2012). There has been an emergence of new models related to sustainability, including models focused on ESG performance, as well as assessments aligned with the United Nations Sustainable Development Goals (SDGs) (Corvo et al., 2021). Various metrics, including cost-benefit analysis, ranking systems, and social accounting and audit, are employed to fulfil the diverse information requirements of different stakeholders. The choice of the metric depends on the specific information needs and objectives of each stakeholder (European Commission, 2015). In the following paragraphs, a selection of some of the most renowned methodologies is presented while Table 1 offers a comparative overview highlighting their main characteristics, strengths, and limitations.

Scientific articles within the nonprofit sector predominantly concentrate on case studies detailing the implementation of Social Return on Investment (SROI). SROI analysis involves calculating the financial and social value of a project, determining the relationship between benefits and costs, and using financial proxies to estimate the social value of non-tradable goods (Öncer, 2019). These types of economic models may not fully capture the nature and full impact of the activities (Antadze and Westley, 2012).

The Balanced Scorecard (BSC) is presented as a performance measurement tool that balances short-term financial goals with long-term strategic objectives and that considers both financial and non-financial measures (Öncer, 2019).

In a Cost-effectiveness analysis (CEA), the costs associated with each intervention are measured and compared against their respective outcomes or benefits. This methodology may not capture intangible impacts (Antadze and Westley, 2012).

Cost-benefit analysis (CBA) is based on monetary valuation and has methodological and conceptual limitations when applied to non-monetary goods (Antadze and Westley, 2012).

The well-known Theory of Change is a framework that explains how activities lead to a series of outcomes,

Table 1. Overview of the key strengths and limitations of the most widely recognized methodologies.

Model	Main focus	Strengths	Limitations
SROI (Social Return on Investment)	Quantifies social value in monetary terms	Provides clear cost-benefit ratio; useful for communicating value in a easily understandable language (money); well-known methodology	Requires financial proxies (skewed proxy when there is no obvious financial value); may oversimplify complex social outcomes; no standardization of the indicators; difficulty in proving the direct effect from the interventions
BSC (Balanced Scorecard)	Performance tracking across financial and non-financial indicators	Balances short-term financial goals with long-term strategic objectives; adaptable	May lack direct impact attribution; No universal or standardized set of indicators
CEA (Cost-Effectiveness Analysis)	Compares costs with specific outcomes	Useful for evaluating intervention efficiency; simple	May not capture qualitative or indirect effects; aggregating cost-effectiveness across multiple outcomes is difficult; sensitive to place, scale, and errors in estimates, it may be difficult to precisely compare programs; no universally standardized indicators
CBA (Cost-Benefit Analysis)	Monetizes all costs and benefits	Offers clear decision-making basis; well-known methodology	Difficult to assign monetary value when applied to non-monetary goods; risk of ignoring social equity; no standardized list of input indicators
Theory of Change	Framework that explains how activities lead to a series of outcomes, ultimately achieving the desired impact	Creative visualization and collective reflection on how to generate social impact; instrument for defining cause and effect relationships	Not a standalone evaluation methodology but a conceptual framework or planning tool

ultimately achieving the desired impact. The flow includes resources, activities, outputs (points of contact with beneficiaries), results (changes achieved in the lives of beneficiaries), and impact (the extent to which results are attributable to the specific activities) (European Commission, 2014b).

The creation of countless models involving measurement processes shows the extent to which there is a clear need to evaluate this social reality, which appears to be complex and multidimensional. An abundance of confusing initiatives has thus contrasted with the need for a simplified and effective measurement system (Barby et al., 2021). The European Commission and many other organizations have started efforts to develop methodologies for measuring SI, but no definitive solution has yet been found (Costa and Pesci, 2016). In other words, despite the multiple efforts, there is a lack of widespread adoption of SIAs in evaluating projects and no methodology has become a universal consensual reference (Burdge, 2002), even though that common logic is clearly needed (Costa and Pesci, 2016).

There is no single measure that is applicable for the whole sector of nonprofit organizations and able to capture changes adequately (Pennerstorfer and Rutherford, 2019). Besides, there is also a lack of clear classification and comparison of these SIA models (Grieco et al., 2015).

First, the lack of a common language makes it difficult to debate and adopt SIA models among professionals (Corvo et al., 2021). There is no universal definition of the concept of SI, with variations in terminology observed between academic fields. The conceptual nature of SI requires flexibility and the consideration of multiple perspectives, stakeholder engagement and context-specific approaches (Lee et al., 2021). Although the term "impact" in the social sector lacks a consistent definition, it is associated with lasting changes in people's lives and it addresses the root causes of social problems (Ebrahim and Rangan, 2014).

Burdge and Vanclay (1996) state that by SI we mean the consequences to human populations of any public or private actions that alter how people live, work, play, relate to one another, organize to meet their needs, and generally act as members of society. Latané states that by SI we mean any of the great variety of changes in physiological states and subjective feelings and emotions, cognitions and beliefs, values, and behavior, that occur in an individual, human or animal, as a result of the real, implied, or imagined presence or actions of other individuals. Freudenburg states that SI refers to impacts (or effects, or consequences) that are likely to be experienced by an equally broad range of social groups as a result of some action. For Gentile, SI is the wider societal concern that reflects and respects the complex interdependency between business practice and society (Maas and Liket, 2011).

The International Association for Impact Assessment (IAIA), considered the leading global authority on impact assessment best practices, defends that SI Assessment are "intended and unintended social consequences, both positive and negative, of planned interventions (policies, programs, plans, projects) and any social change processes invoked by those interventions" (IAIA, n.d.).

On the other hand, the European Commission has a different definition regarding SI: "The reflection of social outcomes as measurements, both long-term and short-term, adjusted for the effects achieved by others (alternative attribution), for effects that would have happened anyway (deadweight), for negative consequences (displacement), and for effects declining over time (drop-off)" (European Commission: Directorate-General for Employment, Social Affairs and Inclusion, 2014).

For Impact Europe (formerly European Venture Philanthropy Association—EVPA), SI is "the attribution of an organization's activities to broader and longer-term outcomes, which are in turn defined as the changes, benefits, learnings, or other effects (positive or negative, both long and short term) that result from an organization's activities" (Impact Europe, n.d.).

On the other hand, Social Value International (SVI) states that Impact is the difference between the outcome for participants, considering "what would have happened anyway, the contribution of others, and the length of time the outcomes last" (Social Value International, 2015).

Regardless of the different perspectives that are observed, there are common elements between them, such as the focus on the long-term changes and on the needs of society with regard to environment, human rights, economic development, education, citizenship, and health (Matzembacher et al., 2019).

Second, the evaluation of social performance is challenging due to the difficulty of establishing causal links between different interventions (Ebrahim et al., 2014).

Third, the interdisciplinary nature of SI measurement: different paradigms and levels of measurement are considered in the academic debate, where there is a clear diversity of beliefs and preferences for measurement methods (Molecke and Pinkse, 2017). Measuring SI must consider the richness and complexity of the construct, including multidimensional approaches. For example, assessing the impact of a youth development program may require indicators that go beyond participation rates or academic improvement. It might also involve capturing changes in self-esteem, sense of belonging, or aspirations for the future—dimensions that are subjective, and evolve over time. Similarly, a community housing initiative may be evaluated not only by the number of homes built, but also by changes in residents' sense of safety, neighborhood cohesion, and access to services. Addressing such complexity and multidimensional reality often requires a combination of quantitative indicators (e.g., changes in income levels or employment) and qualitative methods (e.g., interviews, case studies, or life stories) to truly capture the depth and nuance of social change (Rawhouser et al., 2019).

Fourth, a conceptual challenge, the oversimplification of complex issues: SI is difficult to measure as it is abstract and not easily captured by metrics, involving a complex and abstract phenomenon (Antadze and Westley, 2012). SI results from emotional actions, help, kindness, aimed at protecting the most vulnerable. For this reason, SI is associated with psychological or emotional concepts, such as "well-being," "personal fulfilment," "autonomy," "faith in the goodness of others," and other similar ones. These are concepts seen as "qualitative," in which the generalization of the group is avoided in favor of the centralization of the single individual. There is, therefore, a certain resistance in measuring this "qualitative" concept, for fear of approaching this enormous emotional dimension, often considered not visible, in a reductive way. This resistance is reinforced by the fact that social areas do not commonly interconnect with analytical areas.

Fifth, the challenge of proportionality: in measuring impact there is a challenge of costs that is not compatible with most dimensions of social organizations. Measuring impact requires specialized knowledge (European Commission, 2014b), longitudinal studies and resource allocation that may be beyond the capabilities of operating organizations and funders (Ebrahim and Rangan, 2014). Some social entrepreneurs consider SI as immeasurable and consider the cost and effort of implementing formal methodologies to be unrealistic (Molecke and Pinkse, 2017).

But these challenges are also faced by governments in assessing social value due to limited resources, public funding constraints, and growing social needs (Corvo et al., 2021).

Sixth, the challenge of time: the importance of considering the time intervals in which activities lead to results, especially when results manifest outside the observation period. It is therefore crucial to measure the long-term impacts of activities designed to benefit society (Rawhouser et al., 2019).

Seventh, social organizations frequently adhere to conventional management methods and may have established structures rooted in tradition. Thus, the introduction of SI measurement and evaluation processes can create internal discomfort and resistance within organizations. Even if organizations understand these assessments as enabling them to secure funding, assessments can threaten autonomy, create anxiety, and challenge organizational values.

Eighth, the potential problem of double counting and cumulative effects (Arce-Gomez et al., 2015) when measuring indirect social effects in an evaluation. Certain social effects may be counted multiple times in different parts of the assessment (Zimdars et al., 2018).

Ninth, clear importance is given to comparability and consistency in measuring SI, where the need for a harmonized approach must consider diversity between sectors. Social enterprises struggle to determine whether the numbers represent marginal or breakthrough success due to the absence of a point of comparison (Molecke and Pinkse, 2017).

Tenth, SIA methods must consider a more comprehensive and inclusive measurement, considering all stakeholders throughout the entire measurement process (Costa and Pesci, 2016). Community engagement helps organizations address

problems and opportunities, obtain valuable feedback, improve community relations, and enhance performance in areas of community interest (Epstein and Roy, 2001). Involving citizens in evaluating results leads to reducing conflicts (Arce-Gomez et al., 2015) and to improving their lives (Soler-Gallart and Flecha, 2022). Disadvantages and challenges of participatory approaches include the need for clarification, management of diverse stakeholders, and determination of realistic SI (Arce-Gomez et al., 2015).

Other challenges include limitations in terms of incomplete data and the need to improve access to data (Antadze and Westley, 2012) and the disconnection between metrics and the reality on the ground (Molecke and Pinkse, 2017). Given the diversity of challenges presented, Table 2 provides a consolidated overview by grouping them into thematic categories, allowing for a clearer understanding of the main conceptual, methodological, operational, and contextual barriers to social impact measurement.

Different entities may require different methods tailored to their specific activities, objectives, and impact aspects they need to measure (Maas and Liket, 2011). Impact investors are looking for uniform measurement standards, but a balance between uniformity (ensuring comparability) and relevance is crucial, enabling specific purposes while enabling sectoral analysis.

Social entrepreneurs often rely on a combination of elements from several methodologies, rather than fully committing to a specific formal methodology for measuring SI. The idea of a universal "gold standard" is often criticized (Costa and Pesci, 2016). Thus, social entrepreneurs turn to "bricolage," a combination of material and ideal practices, allowing to overcome the limitations of formal methodologies and be more aligned with their own understandings and experiences (Molecke and Pinkse, 2017).

The variety of models and the absence of a single model that suits all organizations challenge managers and academics. Taking a step behind, classification systems have been proposed to help organizations understand the main characteristics of models. They allow the identification of macro categories of models based on several variables, offering more personalized options for social entrepreneurs (Grieco et al., 2015).

Dimensions such as purpose, deadline, orientation (input or output), duration of time (short or long term), perspective (micro, meso, or macro), approach (process methods, impact methods, monetization methods), and generalization of application (multi-sector vs. single sector) (Rawhouser et al., 2019) are described as factors that differentiate methods. These characteristics influence the purpose, the focus, and the results of measurement. Some methods can assess impacts prospectively, focusing on the expected outcomes of planned programs. On the other hand, Öncer (2019) identifies seven analysis variables for mapping SI assessment models: data typology, impact typology, purpose, model complexity, sector, period, and entity. Studies focus on various aspects of SIA models, including their measurability, purpose, data viability, sectoral applicability, function (e.g., monetization, evaluation, management), and usability of tools to measure SI (Corvo et al., 2021).

Materials and Methods

The work will be segmented into two sections: (1) mapping and classifying SIA models; and (2) mapping the indicators related to these models. In the latter phase, efforts will be made to harmonize indicator nomenclature through a combination of manual and automated methods.

The study type comprises a Literature (Narrative) Review followed by a classification of methodologies and a listing of corresponding indicators.

The databases employed for sourcing methodologies encompassed Google Scholar, Scopus, ResearchGate, and EBSCO Host. The investigation utilized a set of keywords, including "social impact" AND "measurement" OR "assessment' OR "analysis" OR "metrics" OR "analytics" OR "evaluation."

Criteria for inclusion and exclusion of methodologies

Employing a comprehensive literature review, a classification of the existing SIA models was carried out. The inclusion criteria encompassed methodologies utilized by social organizations, private enterprises, and public entities for measuring impact on individuals, thereby excluding assessments on entities and on organizations themselves. It also considered both internal impact (on employees, on customers, and on suppliers) and external impact (on the community).

There were excluded from the mapping methodologies (1) based on formulas or generalized statistical models not specialized in the social sector ($N=45$); (2) focused on ecology and ecosystems, recognizing it as a distinct area warranting a dedicated analysis ($N=35$); (3) for which it was not possible to obtain any information available in the literature ($N=23$); (4) that were repeated (referenced by another name) ($N=22$); (5) related to the financial and governance aspects of companies or organizations ($N=5$); (6) that were purely visual tools ($N=2$); and (7) other reasons (e.g., impact evaluation within the government public sector) ($N=8$). Applying these exclusion criteria, the initial pool of 284 methodologies was narrowed down to 144 (Figure 1). In the **Appendix** chapter, the list of the 144 SIA models is presented (Table A1**)**.

Dimensions of the classification table

Therefore, a classification table was created mapping the following dimensions, for each methodology: Methodology name, Brief explanation, Type of methodology, Formula, Result, Formula details, Result details, Areas, Stakeholders,

Table 2. Classification of key challenges in measuring social impact.

Category	Challenge	Description
Conceptual	Lack of a universal definition of SI	Multiple interpretations make standardization and consensus difficult.
	Qualitative nature of social impact	Emotional and psychological elements like well-being and fulfillment are hard to quantify.
	Oversimplification risk	Reducing complex realities into indicators can lead to loss of depth.
Methodological	Considering all stakeholders	Many models fail to include the full range of stakeholders impacted—such as employees, suppliers, and communities.
	Difficulty establishing causal links	Attributing outcomes directly to specific interventions is often complex.
	Double-counting and cumulative effects	Risk of counting the same impact more than once when measuring indirect effects.
Operational	High cost and effort of measurement	Many methodologies require specialized knowledge, time, and resources that small organizations lack.
	Resistance within organizations	Evaluation processes may create discomfort or fear of accountability among staff.
	Data limitations	Incomplete or inaccessible data can affect the accuracy and credibility of results.
Contextual/External	Sectoral diversity	Different entities (e.g., nonprofits, companies, public sector) have diverse goals and require tailored approaches.
	Time lag of impact	Some social impacts manifest only in the long term, outside observation periods.

No. of fixed indicators, No. of example indicators, Predefined indicators, Type of organization, Strengths, Weaknesses, Impact indicators (example), Observations and Source (Table 3).

The type of methodology is divided into seven categories:

1. **Quantitative**: Methodologies that produce a measurable, numeric outcome, such as a score, ratio, or monetary value.
2. **Qualitative**: Methodologies based primarily on narrative descriptions, case studies, or subjective evaluation without numeric outputs.
3. **Framework**: Conceptual, logical models or guiding structures that outline steps, principles, or structures for impact assessment but do not specify specific metrics (e.g., Theory of Change).
4. **Standard**: Established, often formally endorsed methodologies used for benchmarking or compliance, which may include certification systems or formal reporting guidelines (e.g., GRI).
5. **Lists of Indicators**: Tools that offer a predefined set of indicators, typically without an underlying methodology for how to apply or interpret them.
6. **Libraries**: Collections or repositories of resources (such as databases of metrics or outcomes) meant to support impact evaluation but not forming a standalone methodology.
7. **Not defined**: This category includes methodologies that lacked sufficient documentation in the available literature to clearly determine their process, structure, or intended output. In these cases, the source material described the approach in abstract or promotional terms without operational detail, making classification uncertain.

While overlaps may exist (e.g., a standard that also includes a list of indicators), categorization was based on the **dominant characteristic** as described in the literature source.

Indicators mapping

Simultaneously, a supplementary classification table was compiled containing a mapping of indicators utilized by the methodologies under consideration.

Indicators serve as indispensable tools in measurement, furnishing quantifiable benchmarks that monitor progress and provide insights into the efficacy of initiatives. They guide informed decision-making and facilitate continuous improvement efforts (Pennerstorfer and Rutherford, 2019).

Indicators were excluded according to the following criteria:

(1) **Duplicates within each methodology** (2370): Identical indicators appearing more than once within the same methodology were removed to eliminate redundancy.
(2) **Financial and governance aspects** ($N = 1154$): Indicators whose focus was primarily financial (e.g., profitability) or administrative (e.g., board composition) were excluded, as they do not directly assess social impact.

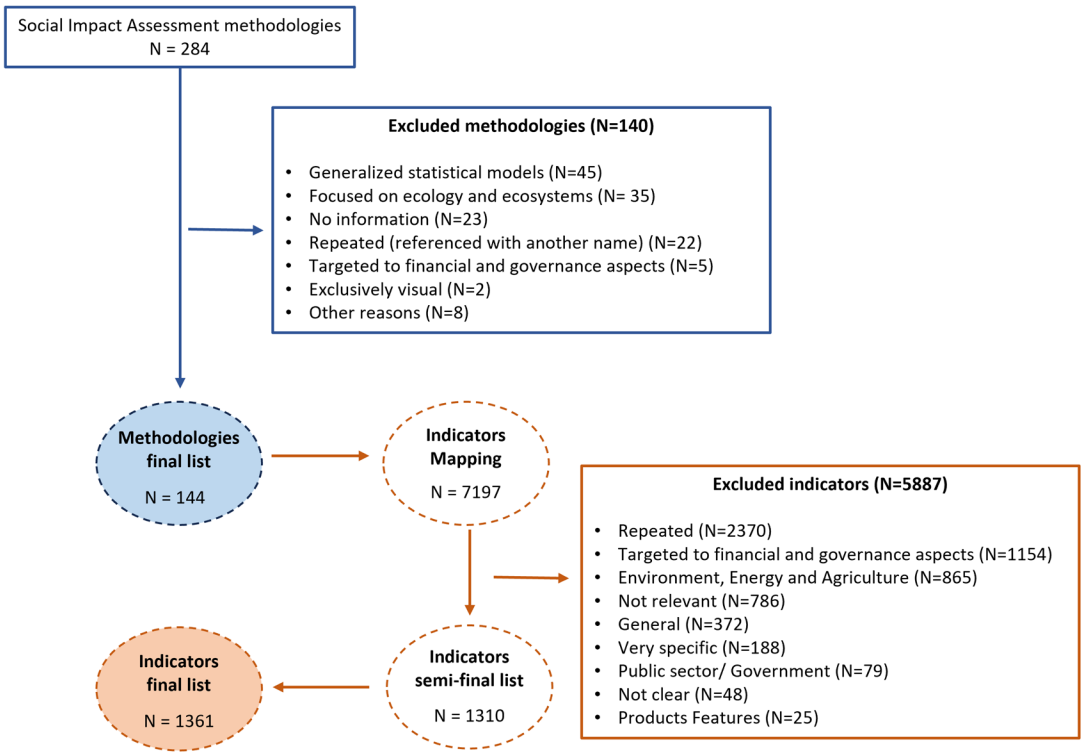

Figure 1. Selection of methodologies and indicators for inclusion in the study, based on the established inclusion and exclusion criteria.

(3) **Environment, Energy and Agriculture** ($N=865$): In alignment with the study's defined scope, indicators addressing environmental sustainability, energy efficiency, or agricultural practices were excluded.
(4) **Not relevant** ($N=786$): Indicators with no clear relation to social outcomes or stakeholder impact—often technical or unrelated to impact measurement—were removed (e.g., "Number of printers installed in the administrative office").
(5) **General** ($N=372$): Indicators considered overly broad or vague, such as "increase impact" were excluded due to a lack of specificity, operational criteria, or defined targets.
(6) **Too specific** ($N=188$): Indicators narrowly tailored to highly contextualized situations (e.g., "Reduction in missed medical appointments among diabetic patients after SMS reminders") were excluded due to their limited applicability across broader assessment frameworks. An exception to removal was made when it was appropriate and feasible to generalize the indicator in a way that preserved its core value, allowing it to be meaningfully included in the final dataset without compromising its relevance.
(7) **Public/government sector focus** ($N=79$): Indicators relevant primarily to public institutions or governance performance were excluded to maintain the focus on civil society and private-sector impact.
(8) **Not clear** ($N=48$): Indicators with ambiguous or poorly defined language, or lacking adequate description in the source material (e.g., "Impact level achieved"), were excluded due to interpretive uncertainty.
(9) **Product characteristics** ($N=25$): Indicators describing only the physical, technical, or performance attributes of products—without direct social implications—were removed.

Applying these exclusion criteria, the initial pool of 7197 indicators was narrowed down to 1310. The final list included 1361 indicators, as some of the 1310 indicators were broken down into multiple indicators (Figure 1).

The mapping included the following identification: Methodology name, Stakeholder, Indicator area, and Indicator name. Only indicators associated with the previously mapped methodologies were included, that is, there were excluded isolated indicators that organizations use in their reports.

Indicator harmonization

After the establishment of the classification table incorporating indicators aligned with the mapped methodologies, a harmonization process ensued. This procedure entailed

Table 3. Classification table structure: dimensions and fields.

Dimensions	Fields
Methodology name	Open
Brief explanation	Open
Type of methodology	Quantitative, Qualitative, Frameworks, Standards, Lists of Indicators, Libraries, and Not defined
Formula	Qualitative, Simple average, Weighted average, Percentage, Cost per Impact, Specific formula (Score), Unknown, and Not applicable
Result	Categories, Checklist, Database, Design Thinking/ Visual, Distribution/Probability, Framework, Indicators, Library, Metrics, Priorities, Qualitative classification, Ranking, Standards/Certification, Others, and Unknown
Formula details	Open
Result details	Open
Areas	Open
Community	Y/N
Environment	Y/N
Organization/company	Y/N
Employees	Y/N
Customers	Y/N
Suppliers	Y/N
No. of fixed indicators	Numeric
No. of example indicators	Numeric
Pre-defined indicators	Y/N
Type of organization	Companies, Organizations, Social Enterprises, Government, Transversal
Strengths	Open
Weaknesses	Open
Impact indicators (example)	Open
Observations	Open
Source	Open

harmonizing the nomenclature and concepts of the numerous indicators under scrutiny. Failure to undertake this harmonization would have stopped progression to the subsequent phase: descriptive analysis.

The harmonization process employed two distinct methodologies: manual and automatic. Manual harmonization confers the advantage of assimilating pre-existing knowledge accrued during the investigation, albeit it remains susceptible to human error and bias. Conversely, the automatic methodology, drawing upon thousands of published databases within the domain, offers the benefit of impartiality. Nevertheless, its reliance on an automated approach may not invariably align with the specificity of this problem. The overarching objective is to place alongside these two methodologies and integrate their outcomes into a unified mapping, facilitating subsequent dimensional aggregation. The results obtained manually and through automatic techniques were compared and compiled into a single harmonization, regarding their name, the stakeholder they refer to, and the impact area they pertain to.

To automate the harmonization of indicators, a Text Mining model was employed.

Specifically, Topic Modeling, a technique within the domain of Natural Language Processing (NLP), was deployed. Topic modeling, an unsupervised learning method, endeavors to discern abstract "topics" or themes within the corpus of text, based on probabilistic latent semantic analysis (Nikolenko et al., 2017). Specifically, the BERT (Bidirectional Encoder Representations from Transformers), the state-of-the-art natural language processing model introduced by Google in 2018, was utilized. To obtain textual representations in a vector space, the BERT model was utilized as a foundational methodological framework. BERT was selected due to its widespread adoption and recognition as a state-of-the-art model for several years. Its ability to generate contextualized representations of textual content provides a more accurate semantic understanding—particularly valuable for analyzing the nuanced and domain-specific language of social impact indicators (Arora et al., 2020).

Subsequently, the Uniform Manifold Approximation and Projection (UMAP) technique was employed for dimensionality reduction, allowing the transformation of high-dimensional textual data into a lower-dimensional space while preserving the underlying structure and relationships. UMAP was employed as it more effectively preserves both local and global data structures. This characteristic ensures that semantically similar indicators remain closely positioned after projection, while dissimilar indicators remain distinct—an essential property for meaningful topic modeling (Wang et al., 2021).

Additionally, Hierarchical Density-Based Spatial Clustering of Applications with Noise (HDBSCAN) served as an advanced clustering algorithm to discern meaningful groupings within the reduced-dimensional space, thereby facilitating the identification of latent patterns and associations inherent within the textual data corpus. HDBSCAN was selected for its high robustness to noise and outliers. These characteristics make it especially well-suited for clustering high-dimensional, unstructured textual data, particularly in contexts where the number of underlying topics is not known in advance (Campello et al., 2013).

BERT, grounded in the Transformer architecture, undergoes pre-training on numerous volumes of text data (3.3 billion words) via unsupervised learning tasks such as Masked Language Modeling (MLM) and next-sentence prediction, alleviating the unidirectionality constraint (Devlin et al., 2018). The Transformer architecture is based on the mechanism of self- attention, which allows the model to weigh the

importance of different words in a sentence during processing (Vaswani et al., 2017). This conceptually simple and empirically powerful bidirectional approach (left-to-right language model, that is, capturing contextual information from preceding and following words in a sentence) (Vaswani et al., 2017) enables it to encapsulate rich contextual information by representing each word within a sentence based on its surrounding context (Devlin et al., 2018).

Specifically, the study utilized the MiniLM model, a refined iteration derived from the BERT model, to enhance its ability in identifying semantically similar sentences. Developed by Microsoft, MiniLM employs various techniques to reduce the number of parameters compared to BERT, thereby enhancing its scalability and enabling fast performance across resource-constrained environments (Wang et al., 2020).

As previous referred, two sub-models were instrumental in this endeavor:

1. UMAP (Uniform Manifold Approximation and Projection): a novel developed method for dimensionality reduction that exhibits nonlinear properties. Its primary objective is to reduce high-dimensional datasets into lower-dimensional representations while retaining both the local and global structural characteristics of the original data (McInnes et al., 2018). UMAP distinguishes itself through its commendable computational efficiency, allowing for swift execution even with substantial datasets. Moreover, it offers a degree of adaptability through Parameter Tuning, enabling users to adjust the balance between preserving local and global structures according to specific requirements. Notably, UMAP demonstrates robust scalability and adeptness in managing extensive datasets, owing to its implementation of an approximate nearest neighbor search algorithm, a feature that facilitates the handling of datasets comprising millions of data points.
2. HDBSCAN (Hierarchical Density-Based Spatial Clustering of Applications with Noise): a density-based clustering algorithm tailored for high-dimensional data. HDBSCAN constructs a hierarchy of clusters based on varying density levels, subsequently identifying clusters with the highest stability as the final clusters (Malzer and Baum, 2020). This methodology is particularly suited for datasets characterized by irregular shapes and varying densities (Asyaky and Mandala, 2021). HDBSCAN utilizes the following parameters:

• min_samples = the minimum number of neighbours requisite for a core point, with higher values leading to the exclusion of more points as noise/outliers (DBScan component of HDBScan). Hence, HDBSCAN is able to identify clusters of points that are surrounded by a large amount of noise or outliers.

• min_cluster_size = minimum size a final cluster can attain, with larger values resulting in larger clusters (H component of HDBScan).

Several values for these parameters were tested, ultimately leading to the decision to create one model with a low value (5) and another with a high value (50), integrating both sets of results.

The employment of dimensionality reduction and clustering models assumes paramount significance in harmonizing the extensive array of analyzed indicators, facilitating the reduction of the original data's dimensionality. Moreover, the grouping of analogous concepts facilitates the translation of apparently disparate concepts sharing common ideas.

The classification tables were created in Microsoft Excel. The automated Harmonization methods were developed in Python version 3.11. The figures were generated using Excel (Figure 2), Python (Figure 3), Venngage (Figures 4–7), and Excalidraw (Figure A1).

The original data presented in the study (both Python Code and Methodologies and Indicators datasets) are openly available in https://github.com/ [The full link was not provided to maintain the anonymity of the article, but the information was included in a different file submitted.].

Results and discussion

Examples of the most well-known and analyzed methodologies in the literature include the Social Return on Investment (SROI), the London Benchmarking Group (LBG), the B Impact Assessment, the Balanced Scorecard (BSc), the Social Costs-Benefit Analysis (SCBA), and the Robin Hood Foundation Benefit-Cost Ratio. The full list of the analyzed methodologies is presented in the Appendix (Table A1).

Methodologies classification: Quantitative dimension, comparability and exhaustiveness

284 methodologies were mapped with the previously identified dimensions. 140 methodologies were excluded based on the exclusion criteria. As a final list, 144 methodologies were mapped. The SIA models classification allowed to obtain 7197 indicators (Figure 1). To our current knowledge, Corvo et al. (2021) (Corvo et al., 2021) have analyzed the largest number of methodologies in a single article, examining 98 methodologies. This record is followed by Grieco et al. (2015) (Grieco et al., 2015), who analyzed 76 methodologies. Therefore, this study surpasses the number of methodologies mapped and compared in scientific articles, totaling 144 methodologies.

A total of 74 out of the 144 methodologies (51%) are considered Quantitative in nature. The remaining methodologies were distributed as being Qualitative ($N=9$), Frameworks ($N=36$), Standards ($N=12$), Lists of indicators ($N=10$),

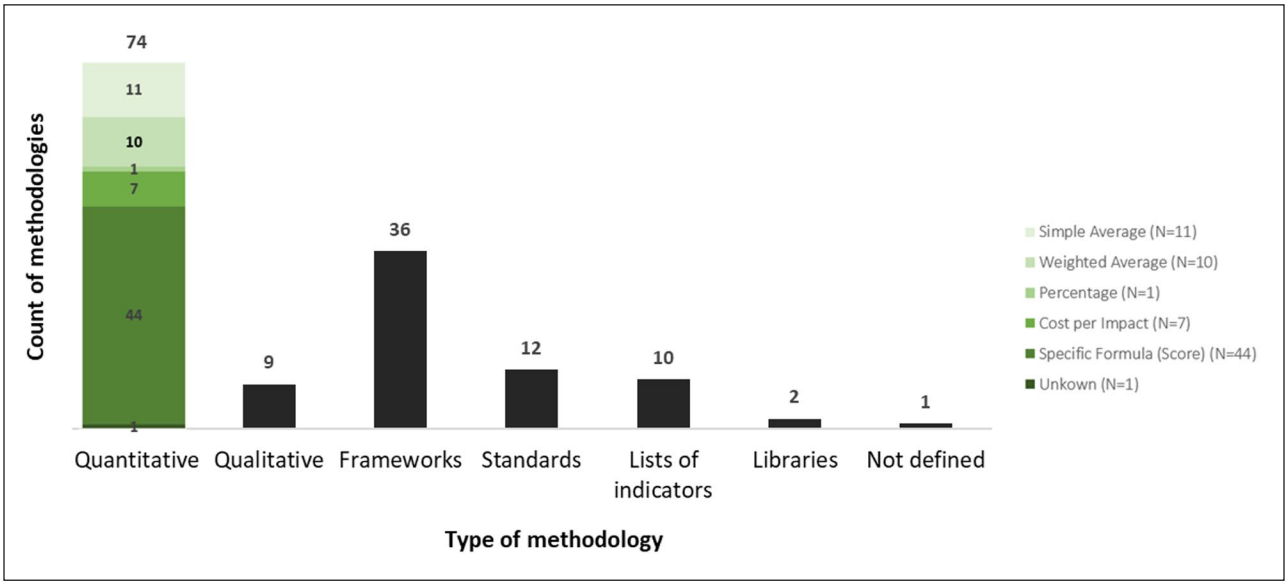

Figure 2. Distribution of methodologies, by type.

Libraries ($N=2$), or Not defined ($N=1$). With regard to the 74 quantitative methodologies, the most frequent calculation method was the Specific Formula ($N=44$), followed by the Simple Average ($N=11$) and the Weighted Average ($N=10$).

As a specific formula, it refers to a metric that is obtained through a set of mathematical operations that are not typically used such as average or percentage (Figure 2).

Only 42 of the 144 (29%) methodologies are aimed to all stakeholders, while the rest are only aimed to some of them (such as, for example, employees and customers or the community in general). 125 (87%) of the methodologies are applicable to the community in general, 77 (53%) measure the impact on organizations or companies globally, 105 (73%) measure the impact on employees, and 67 (47%) measure the impact on customers or suppliers (Figure 3).

Of the 74 quantitative methodologies, 15 were exhaustive, that is, meeting all stakeholders. Of these 15, only 4 include pre-defined indicators, a fundamental characteristic for comparability. However, 3 of these 4 are only applicable to companies and not to non-profit organizations. The remaining methodology does not allow applicability to any area of intervention: it only includes indicators relating to the economic growth and innovation of organizations.

Thus, no SIA model was found that simultaneously exhibits the three fundamental characteristics: quantitative nature, exhaustiveness, and comparability.

Indicators mapping and harmonization

Regarding the analysis at indicator level, of the 7197 indicators initially listed, 5887 were excluded. Of the 1310 indicators that remained from the screening according to the exclusion criteria, some were divided into multiple ones

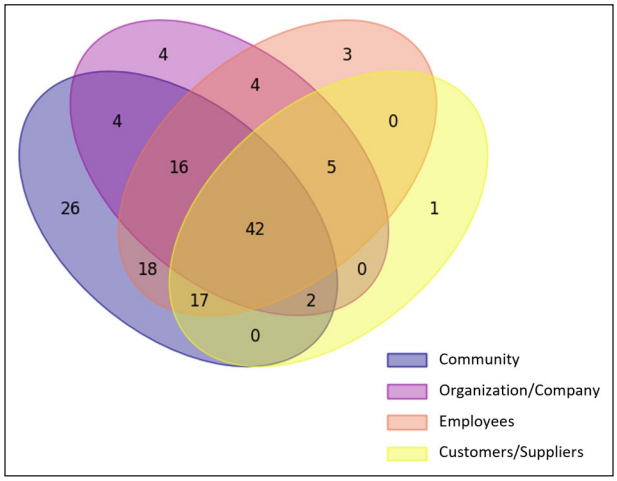

Figure 3. Distribution of methodologies based on the stakeholders they target.

when they represented different concepts, resulting in a final list of 1361 indicators (Figure 1).

The indicators underwent harmonization regarding their nomenclature, initially through manual methods and subsequently via Text Mining techniques, as delineated in the methods section. The results obtained manually and through automatic techniques were compared and compiled into a single harmonization, regarding their name, the stakeholder they refer to, and the impact area they pertain to.

The 1361 indicators from the final list were distributed according to the stakeholder they represented and their impact areas, having obtained the distribution shown in Figures 4 through 7. The size of the circles represents the proportion of indicators that each area includes, compared

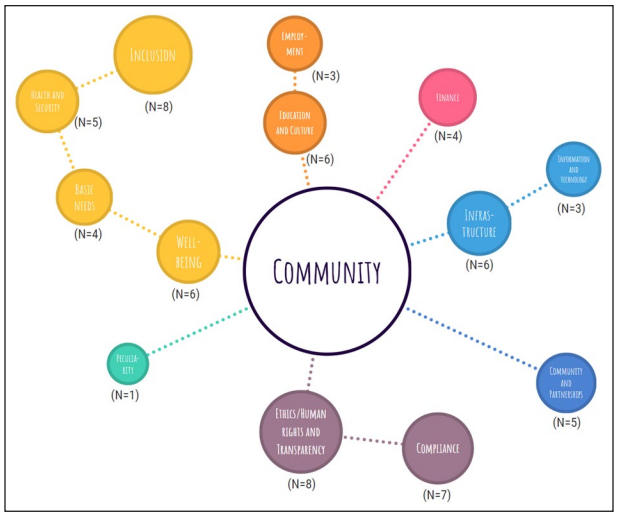

Figure 4. Distribution of indicator areas concerning the stakeholder **community**.

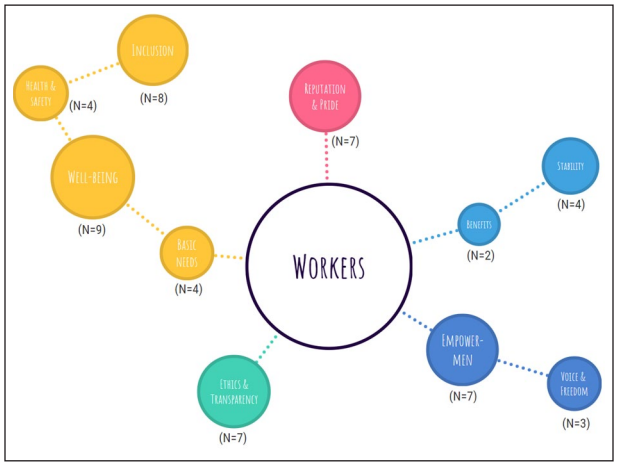

Figure 5. Distribution of indicator areas concerning the stakeholder **workers**.

to the others. The number of indicators presented (N) in each circle refer to the count of unique indicators within each area and each stakeholder. They totalize 153 unique indicators.

Conclusion and future directions

Since the onset of the 21st century, the discussion regarding social responsibility has become a central focus not only for social organizations but also within the corporate sector (Corvo et al., 2021). The creation of numerous models for measuring impact underscores the clear necessity to evaluate this complex and multidimensional social reality (Barby et al., 2021).

A universal measure that comprehensively captures changes across the nonprofit sector remains elusive (Burdge,

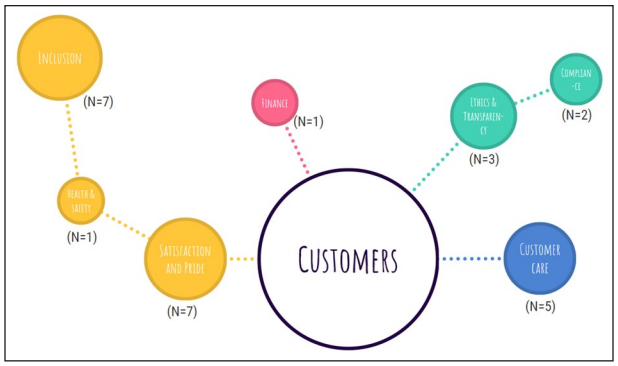

Figure 6. Distribution of indicator areas concerning the stakeholder **customers**.

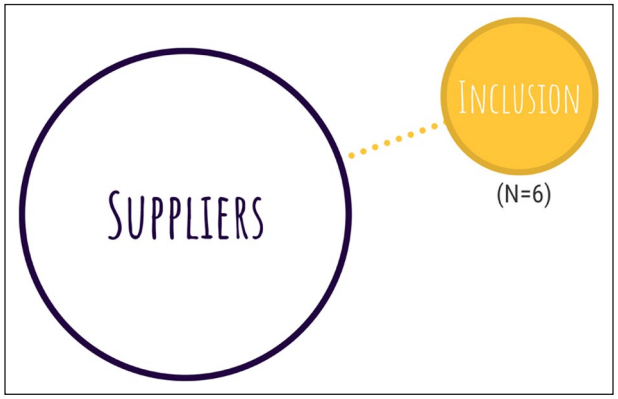

Figure 7. Distribution of indicator areas concerning the stakeholder **suppliers**.

2002; Pennerstorfer and Rutherford, 2019). Moreover, there exists a deficiency in the clear classification of these models for assessing SI (Grieco et al., 2015).

The literature outlines numerous challenges associated with measuring SI. First, a historical reason: there has not been a progressive integration of the strengths of each model into a single one, but rather a new methodology independent of the previous ones has always been created, which does not allow for the evolution and robustness of a single model. Second, institutions of transnational responsibility (such as the European Commission) have difficulty in adopting an official measurement or even supporting the creation of a reference in this area (Costa and Pesci, 2016). In part, this difficulty arises from the lack of a consensual definition of SI (Corvo et al., 2021). Without this prior harmonization work, it is not possible to progress towards a single model. It will be very important to accept a definition for the concept of SI that is as comprehensive from a theoretical point of view as it is based on concrete and specific characteristics of each reality. Third, a conceptual reason: SI results from emotional actions, help, kindness, aimed at protecting the most vulnerable. SI is associated with psychological or emotional concepts, such as "well-being," "personal fulfilment," "autonomy," "faith in

the goodness of others," and other similar ones. These are concepts seen as "qualitative," in which the generalization of the group is avoided in favor of the centralization of the single individual. There is, therefore, a certain resistance in measuring this "qualitative" concept, for fear of approaching this enormous emotional dimension, often considered not visible, in a reductive way (Antadze and Westley, 2012). Fourth, different entities may require different methods tailored to their specific activities, objectives, and impact aspects they need to measure. Social entrepreneurs often rely on a combination of elements from several methodologies, rather than fully committing to a specific formal methodology for measuring SI (Costa and Pesci, 2016).

These challenges give rise to significant concerns: on one hand, the importance of considering all stakeholders, realizing that SI does not result only from specific actions from social sector entities within the community, it is also a consequence of the daily activity of institutions on employees, suppliers, and customers (Exhaustiveness) (Costa and Pesci, 2016; Miller et al., 2007). On the other hand, the importance of using analytical knowledge to transform emotional concepts into indicators that can be observed and, consequently, into quantitative variables (Quantitative dimension) (Grieco et al., 2015). Also the importance of measuring every dimension, otherwise there is no comparability (in space, in time, by type of promoting entity, by type of action, by type of beneficiary) (Molecke and Pinkse, 2017). Lastly, the importance of creating observation and collection methods that can be relatively quick and financially accessible, making the model more operational and capable of being applied by a maximum number of users (Proportionality) (European Commission, 2014b). To this end, it is very important to know the statistical indicators available, nationally and internationally, to avoid specific collections through surveys, although these must always exist. Some common sense is also necessary in the decision regarding the total number of indicators because although the definition of SI is multidimensional, especially if we consider it applied to several stakeholders, the number of indicators for each subdimension identified must be reduced to a minimum.

Given the challenges outlined, there is a clear and pressing need: developing robust classification systems that enable an understanding of the strengths and weaknesses of each model, and help professionals select the methodology that best aligns with their specific needs. Given this pressing need, our study first focused on this classification. Our analysis revealed that no SIA model simultaneously presents the three fundamental characteristics: quantitative nature, exhaustiveness, and comparability. This prompts inquiry into whether the absence of a universal reference stems from a pragmatic perspective, necessitating methodologies tailored to individualized contexts, or from the continued absence of a comprehensive model that integrates all fundamental attributes.

As a prospective path for further research, the proposition persists in interrogating the preceding inquiry. Should it be discerned that the absence of a universal methodology stems from the absence of one possessing the delineated attributes, there arises a pressing necessity to attempt toward its creation. Interviewing organizations spanning diverse maturity levels and operational domains would be instrumental in elucidating this inquiry.

This study presents as its main limitations, on one hand, the absence of the environmental component. This was deemed as an exceedingly comprehensive, multidimensional, and specific domain warranting a dedicated investigation. On the other hand, the exclusion of methodologies assessing the impact on organizational structure (but on individuals) may render this study somewhat restrictive. Nevertheless, once again, the intricacy and specificity of this domain necessitates an independent investigation.

This study represents an advancement in academic practice as it, on one hand, encompasses the most extensive array of models documented in any single academic work to date. On the other hand, it delineates fundamental characteristics for an impact measurement methodology and identifies those characteristics that, while highly significant, are absent from any existing methodology concurrently.

In summary, this study not only presents a comprehensive analysis of diverse methodologies for measuring SI but also underscores the imperative for future research to integrate essential yet currently absent criteria, thereby enhancing the efficacy and applicability of impact measurement frameworks in addressing complex societal challenges.

Authors contribution

Matilde Valente Rosa Main author, who developed the entire study and analysis and responsible for writing the article
Leonor Bacelar-Nicolau Review and monitoring of the study, contribution with literature review (especially in the field of **modelling** and **data analysis**) and article review
Maria de Fátima Ferreiro Review and monitoring of the study, contribution with literature review (especially in the field of **Political Economy**) and article review

Acknowledgements

Acknowledgments are due to João Guarda for his consistent kindness and unwavering support throughout this journey, particularly his invaluable assistance with text mining techniques.

Declaration of conflicting interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

Ethical statement

Not applicable

Consent to participate

Not applicable

Consent for publication

Not applicable

ORCID iD

Matilde Valente Rosa https://orcid.org/0000-0001-8310-0303

Data availability statement

The original data presented in the study (both Python Code and Methodologies and Indicators datasets) are openly available in https://github.com/MatildeVRosa/Social-Impact.

References

Antadze N and Westley FR (2012) Impact metrics for social innovation: Barriers or bridges to radical change? *Journal of Social Entrepreneurship* 3(2): 133–150.

Arce-Gomez A, Donovan JD and Bedggood RE (2015) Social impact assessments: Developing a consolidated conceptual framework. *Environmental Impact Assessment Review* 50: 85–94.

Arora S, May A, Zhang J, et al. (2020) Contextual embeddings: When are they worth it? In: *Proceedings of the 58th annual meeting of the Association for Computational Linguistics*, Online (virtual), 5–10 July 2020. Association for Computational Linguistics.

Arvidson M and Lyon F (2014) Social impact measurement and non-profit organisations: Compliance, resistance, and promotion. *VOLUNTAS: International Journal of Voluntary and Nonprofit Organizations* 25(4): 869–886.

Asyaky MS and Mandala R (2021) Improving the performance of HDBSCAN on short text clustering by using word embedding and UMAP. In: *2021 8th international conference on advanced informatics: Concepts, theory and applications (ICAICTA)*, Bandung, Indonesia, 29–30 September 2021, pp. 1–6. New York, NY: IEEE.

Barby C, Barker R, Cohen R, et al. (2021). Measuring purpose: An integrated framework. *SSRN Electronic Journal [Preprint]*. DOI: 10.2139/ssrn.3771892.

Bengo I, Boni L and Sancino A (2022) EU financial regulations and social impact measurement practices: A comprehensive framework on finance for sustainable development. *Corporate Social Responsibility and Environmental Management* 29(4): 809–819.

Bernardino S and Santos J (2020) Medição de resultados e desempenho organizacional: um estudo exploratório às perceções dos responsáveis de organizações sociais em Portugal. *Portuguese Journal of Finance, Management and Accounting* 6(12): 37–60.

Burdge RJ (2002) Why is social impact assessment the orphan of the assessment process? *Impact Assessment and Project Appraisal* 20(1): 3–9.

Campello RJGB, Moulavi D and Sander J (2013) Density-based clustering based on hierarchical density estimates. In: Pei J, Tseng VS, Cao L, et al. *Advances in Knowledge Discovery and Data Mining. PAKDD 2013. Lecture Notes in Computer Science*. Berlin, Heidelberg: Springer, pp.160–172.

Carroll AB (2009) A history of corporate social responsibility. In: Crane A (ed.) *The Oxford Handbook of Corporate Social Responsibility*. Oxford: Oxford University Press, pp.19–46.

Corvo L, Pastore L, Manti A, et al. (2021) Mapping social impact assessment models: A literature overview for a future research agenda. *Sustainability* 13(9): 4750.

Costa E and Pesci C (2016) Social impact measurement: Why do stakeholders matter? *Sustainability Accounting, Management and Policy Journal* 7(1): 99–124.

Devlin J, Chang M-W, Lee K, et al. (2018) BERT: Pre-training of deep bidirectional transformers for language understanding. In: *Proceedings of NAACL-HLT 2019*, Minneapolis, Minnesota, 2–7 June 2019. Stroudsburg, PA: ACL.

Ebrahim A, Battilana J and Mair J (2014) The governance of social enterprises: Mission drift and accountability challenges in hybrid organizations. *Research in Organizational Behavior* 34: 81–100.

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

Epstein MJ and Roy M-J (2001) Sustainability in action: Identifying and measuring the key performance drivers. *Long Range Planning* 34(5): 585–604.

Esteves AM, Franks D and Vanclay F (2012). Social impact assessment: The state of the art. *Impact Assessment and Project Appraisal* 30(1): 34–42.

European Commission (2014a) Disclosure of non-financial information by certain large companies: European Parliament and Council reach agreement on Commission proposal to improve transparency. Available at: http://europa.eu/rapid/press-release_STATEMENT-14-29_en.htm (accessed 9 June 2023).

European Commission (2014b) *Proposed approaches to social impact measurement*. Brussels: European Commission. Available at: https://social-economy-gateway.ec.europa.eu/document/download/c1bd2c6d-f51c-43d8-b74b-b02ead31b01b_en?filename=proposed+approaches+to+social+impact+measurement+in-KE0414665ENN_0.pdf

European Commission (2015)*Policy brief on social impact measurement for social enterprises*. Brussels: Publications Office of the European Union. Available at: https://op.europa.eu/en/publication-detail/-/publication/19c3e101-f673-437f-a9fe-4a6dc7ff1f6e

European Commission: Directorate-General for Employment, Social Affairs and Inclusion (2014) *Proposed Approaches to Social Impact Measurement in European Commission Legislation and in Practice Relating to EuSEFs and the EaSI*. Brussels, Belgium: Eu Publications.

Grieco C, Michelini L and Iasevoli G (2015) Measuring value creation in social enterprises. *Nonprofit and Voluntary Sector Quarterly* 44(6): 1173–1193.

Impact Europe (n.d.) Impact Glossary - Social impact definition https://www.impacteurope.net/impact-glossary#:~:text=all%20investment%20decisions.-,Social%20impact,result%20from%20an%20organisation's%20activities (accessed 10 April 2024).

International Association for Impact Assessment (IAIA) (n.d.) *Social impact assessment definition*. Available at: https://www.iaia.org/wiki-details.php?ID=23 (accessed 10 April 2024).

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

Lall SA (2019) From legitimacy to learning: How impact measurement perceptions and practices evolve in social enterprise–social finance organization relationships. *VOLUNTAS: International Journal of Voluntary and Nonprofit Organizations* 30(3): 562–577.

Lee EKM, Lee H, Kee CH, et al. (2021) Social impact measurement in incremental social innovation. *Journal of Social Entrepreneurship* 12(1): 69–86.

LeRoux K and Wright NS (2010) Does performance measurement improve strategic decision making? Findings from a national survey of nonprofit social service agencies. *Nonprofit and Voluntary Sector Quarterly* 39(4): 571–587.

Liu M, Liu L and Feng A (2024) The impact of green innovation on corporate performance: An analysis based on substantive and strategic green innovations. *Sustainability* 16(6): 2588.

Maas K and Liket K (2011) Social impact measurement: Classification of methods. In: Burritt R, Schaltegger S, Bennett M, et al. (eds) *Environmental Management Accounting and Supply Chain Management. Eco-Efficiency in Industry and Science*. Dordrecht: Springer, pp.171–202.

Malzer C and Baum M (2020) A hybrid approach to hierarchical density-based cluster selection. In: *2020 IEEE international conference on multisensor fusion and integration for intelligent systems (MFI)*, Karlsruhe, Germany, 14–16 September 2020, pp.223–228. New York, NY: IEEE.

Matzembacher DE, Raudsaar M, de Barcellos MD, et al. (2019). Sustainable entrepreneurial process: From idea generation to impact measurement. *Sustainability* 11(21): 5892.

McInnes L, Healy J and Melville J (2018). UMAP: Uniform manifold approximation and projection for dimension reduction. *Journal of Open Source Software* 3(29): 861.

Miller E, Buys L and Summerville J (2007) Quantifying the social dimension of triple bottom line: Development of a framework and indicators to assess the social impact of organisations. *International Journal of Business Governance and Ethics* 3(3): 223.

Molecke G and Pinkse J (2017) Accountability for social impact: A bricolage perspective on impact measurement in social enterprises. *Journal of Business Venturing* 32(5): 550–568.

Nikolenko SI, Koltcov S and Koltsova O (2017) Topic modelling for qualitative studies. *Journal of Information Science* 43(1): 88–102.

Öncer AZ (2019) Performance measurement in social enterprises. In: Iyigun N (ed.) *Creating Business Value and Competitive Advantage With Social Entrepreneurship*. Hershey, PA: IGI Global Scientific Publishing, pp.205–231.

Pennerstorfer A and Rutherford AC (2019) Measuring growth of the nonprofit sector: The choice of indicator matters. *Nonprofit and Voluntary Sector Quarterly* 48(2): 440–456.

Peris-Ortiz M, Ferreira JJ, Farinha L, et al. (2016) Introduction to multiple helix ecosystems for sustainable competitiveness. In: Peris-Ortiz M, Ferreira J, Farinha L, et al. (eds) *Multiple Helix Ecosystems for Sustainable Competitiveness. Innovation, Technology, and Knowledge Management*. Chamin: Springer, pp.1–13.

Rawhouser H, Cummings M and Newbert SL (2019) Social impact measurement: Current approaches and future directions for social entrepreneurship research. *Entrepreneurship Theory and Practice* 43(1): 82–115.

Social Impact Assessment definition (n.d.) Available at: https://www.iaia.org/wiki-details.php?ID=23 (accessed 10 April 2024).

Social Value International (2015) Discussion document on the valuation of social outcomes. Available at: https://socialvalueuk.org/wp-content/uploads/2017/09/Valuation-of-Social-Outcomes-pdf-1.pdf (accessed 10 April 2024).

Soler-Gallart M and Flecha R (2022) Researchers' perceptions about methodological innovations in research oriented to social impact: Citizen evaluation of social impact. *International Journal of Qualitative Methods* 21: 160940692110676.

Vaswani A, Shazeer N, Parmar N, et al. (2017) Attention is all you need. In: *NIPS'17: proceedings of the 31st international conference on neural information processing systems*, Long Beach, CA, 4–9 December 2017. New York, NY: IEEE.

Wang W, Wei F, Dong L, et al. (2020) MiniLM: Deep self-attention distillation for task-agnostic compression of pre-trained transformers. In: *34th conference on neural information processing systems (NeurIPS 2020)*, Vancouver, Canada, 6–12 December 2020. Advances in Neural Information Processing Systems (NeurIPS).

Wang Y, Huang H, Rudin C, et al. (2021) Understanding how dimension reduction tools work: An empirical approach to deciphering t-SNE, UMAP, TriMAP, and PaCMAP for data visualization. *Journal of Machine Learning Research* 22: 1–73.

Weber M (2008) The business case for corporate social responsibility: A company-level measurement approach for CSR. *European Management Journal* 26(4): 247–261.

Zimdars C, Haas A and Pfister S (2018) Enhancing comprehensive measurement of social impacts in S-LCA by including environmental and economic aspects. *The International Journal of Life Cycle Assessment* 23(1): 133–146.

Author biographies

Matilde Valente Rosa Graduated in Applied Mathematics and Computation and with a Master's degree in Statistics (University of Bristol). Current PhD student in Information Sciences and Technologies. Invited Professor at Universidade de Lisboa. In addition to her role on the Board of Directors at a NGO, she coordinated numerous projects aimed at promoting social integration in Africa and Portugal. Founder of a company specialized in Statistics and Data analysis, Business Intelligence, and Social Impact measurement.

Leonor Bacelar-Nicolau Professor and researcher at Católica Medical School and CIIS-Center for Interdiciplinary Research in Health; founding partner of DataScience Cons; collaborator at ISAMB-Institute of Environmental Health of Faculdade de Medicina of Universidade de Lisboa. Teaching and researching areas include multivariate statistics, data science, health economics and evaluation, health impact assessment, epidemiology, public health and preventive medicine, and infodemiology. Independent expert evaluator for the European Commission.

Maria de Fátima Ferreiro Full professor at the Department of Political Economy (Iscte- Lisbon University Institute). Integrated researcher at CEI-Iscte-Lisbon University Institute. Teaching and research activities in the areas of economic ideas, sustainability, solidarity economics, and territory. Recent international and national funded research projects and publications in the areas of social innovation, food security, and territory.

Appendix A

Table A1. List of the analysed methodologies.

No.	SIA model
1	AA1000
2	Actionable Impact Management (AIM)
3	Acumen Fund's BACO Ratio (Best Available Charitable Option)
4	Aeris Ratings
5	Anticipated Impact Measurement and Monitoring (AIMM)
6	Atkinsson Compass Index of Sustainability
7	B Impact Assessment (BIA)
8	Balance Model
9	Balanced Scorecard (BSC)
10	Bayesian Impact Evaluation
11	BoP Impact Assessment Framework
12	Bridges Ventures Impact Radar
13	Business Ethics Excellence (BEE) Model
14	Business for Societal Impact (B4SI)
15	C3Perform
16	Center for High Impact Philanthropy Cost per Impact
17	CERISE Social Business Scorecard
18	CERISE Social Performance Indicators (SPI)
19	CERISE-IDIA (Impact-Driven Investor Assessment)
20	Charity Assessment Method of Performance (CHAMP)
21	Community Impact Mapping
22	Constituent Voice™ Methodology
23	Cooperative Performance Indicator (CPI) tool
24	Corporate Governance Toolkit
25	Cost-benefit analysis (CBA)
26	Cost-effectiveness analysis (CEA)
27	Cradle-to-cradle certification
28	DALBERG APPROACH
29	DCED Standard for Results Measurement
30	Disability-Adjusted Life Years (DALYs)
31	EFQM Diagnostic Tool: RADAR
32	ESG Disclosure score
33	ESG Risk Rating
34	Fit for purpose
35	Fitch Ratings ESG Relevance Scores
36	FTSE ESG Index
37	Gender Impact Assessment
38	Genuine Progress Indicator (GPI)
39	Global Impact Investing Rating System (GIIRS)
40	Global Reporting Initiative (GRI)
41	Good Jobs Scorecard
42	HERO Scorecard
43	HIPSO Harmonized Indicators for Private Sector Operations
44	Human Development Index (HDI)
45	Human Impact + Profit (HIP™) scorecard
46	IFC' DOTS (Development Outcome Tracking System)
47	Impact Analysis for Corporate Finance & Investments (Tool prototype)
48	Impact Due Diligence Tools
49	Impact Management Project
50	Impact Multiple of Money (IMM)
51	Impact Radar
52	Impact Risk Classification (IRC)
53	Impact Weighted Accounts (IWA)
54	Integrated Questionnaire for the Measurement of Social Capital (SC-IQ)
55	Investors in people (IiP)
56	IRIS+ - Impact Reporting and Investment Standards
57	IRIS+ Impact Toolkit [Measurement technique]
58	ISO 9001
59	ISS ESG Corporate Rating
60	ISS SDG Impact Rating
61	Kirkpatrick's Four-Level Training Evaluation Model
62	Lean Data
63	Life cycle sustainability assessment (LCSA)
64	LOCAL MULTIPLIER 3 (LM3)
65	Logic Model Framework
66	LuxFLAG ESG Label
67	MDG Scan online
68	Measuring the Wellbeing Economy
69	Methodology For Impact Analysis And Assessment (MIAA)
70	MetODD-SDG
71	MFI Factsheet
72	MicroRate Social Rating
73	Lewin's three-stage model of change
74	MSCI ACWI Sustainable Impact Index
75	MSCI ESG Ratings Methodology
76	Multi-Attribute Utility Theory (MAUT)
77	Multi-Criteria Appraisal (MCA)
78	NPC's Charity Analysis Framework
79	Ongoing Assessment of Social Impacts (OASIS)
80	Outcome Mapping
81	Outcome star
82	Participatory Impact Assessment (PIA)
83	Poverty and Social Impact Analysis (PSIA)
84	Poverty Probability Index (PPI®) / Progress out of Poverty Index (PPI)
85	Poverty Stoplight
86	Practical quality assurance system for small organisations (PQASSO)
87	Product SIA
88	Prove It! Toolkit
89	Public Value Scorecard (PVSC)
90	Quality-Adjusted Life Years (QALYs)
91	Questant Process

Table A1. (Continued)

No.	SIA model
92	RISE (Real Indicators Of Success In Employment)
93	Robeco's SDG Framework
94	Robin Hood Foundation Benefit-Cost Ratio
95	S&P Dow Jones Indices and RobecoSAM (Sustainable Asset Management)
96	SASB Standard - Sustainability Accounting Standards Board
97	SDG Compass (Inventory of Business Indicators)
98	SDG Impact Practice Standards
99	SDG Impact Standards
100	Shared Impact Assessment and Measurement Toolbox (SIAMT)
101	Shared Value Measurement
102	SIA (SIA)
103	SIM tool survey
104	Social accounting and auditing (SAA)
105	Social Accounting Matrix (SAM)
106	Social Compatibility Analysis (SCA)
107	Social Enterprise Balanced Scorecard (SEBS)
108	Social enterprise mark
109	Social e-valuatorTM
110	Social Footprint
111	Social IMPact measurement for Local Economies (SIMPLE)
112	Social Impact Quotient (SIQ)
113	Social Life Cycle Assessment (S-LCA)
114	Social multi-criteria evaluation (SMCE)
115	Social Progress Index (SPI)
116	Social rating
117	Social Return Assessment (SRA)
118	Social Return on Investment (SROI)
119	Social Risk Factors (SRF)
120	Social Value Maturity Index
121	Social Value Metrics
122	Socio-Economic Assessment Toolbox (SEAT)
123	Soft outcome universal learning (SOUL)
124	SRI (Socially Responsible Investment) label
125	Stakeholder Value Added (SVA)
126	Standard Ethics Rating (SER)
127	Star Social Firm Quality Standard
128	Success Measures Data System (SMDS)
129	Sustainable Livelihoods Approach
130	Sustainable Performance Index (SPI)
131	The Committee on Sustainability Assessment (COSA) Methodology
132	The Hewlett Foundation's Expected Return
133	The SCALERS Model
134	The values based checklist for Social Firms
135	Theory of Change (ToC)
136	Toniic Impact Portfolio Tool
137	Tool for Indicator Design
138	Toolbox for Analyzing Sustainable Ventures in Developing Countries
139	United Nations Sustainable Development Goals (SDGs)
140	Volunteering Impact Assessment (VIA)
141	Wallace Self-Assessment Tool
142	WBCSD Measuring Impact Framework
143	Wellbeing at work
144	Y Analytics

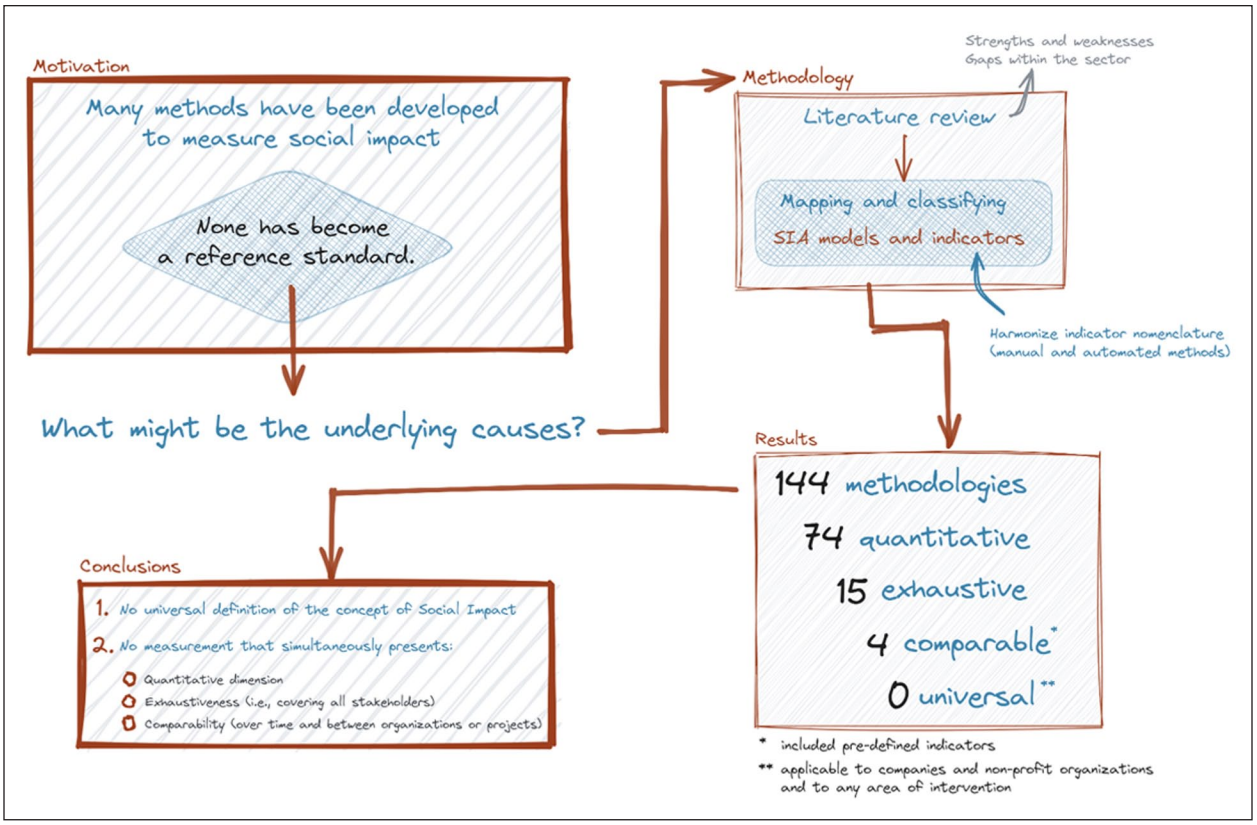

Figure A1. Graphical abstract resuming the study.