Small Bus Econ
https://doi.org/10.1007/s11187-025-01101-7

RESEARCH ARTICLE

From ambition to evidence: a practical tool for startup impact assessment

Christin Eckerle · Orestis Terzidis

Accepted: 14 July 2025

Abstract Impact-driven startups face major challenges in measuring and reporting their social and environmental impacts. In a related manner, impact investors face challenges when attempting to assess the impact of startups. This is primarily due to a lack of comparability and transparency. Despite the proliferation of tools and frameworks, a misfit persists between the capabilities of startups and the expectations of investors. Resource constraints have a significant impact on the capacity of startups to engage in complex impact measurement. Current impact assessment tools are often ill-suited to the needs of startups, as they require extensive data that smaller companies may not have readily available. The present study proposes a design science research approach to derive design knowledge for the impact assessment of startups and subsequently develop a new tool tailored to the needs of startups. The purpose of this tool is to enable startups to measure and report their potential impact efficiently. The objective is to enhance transparency, comparability, and legitimacy, thereby helping startups attract investment while aligning their business activities with their social and environmental missions.

Plain English Summary Impact-driven startups encounter difficulties in measuring their impact due to resource constraints and the complexity of available tools. A new assessment tool will help startups to boost transparency and attract impact investment. This study explores how impact-driven startups currently measure their social and environmental impacts, and the challenges they face in doing so due to the complexity and resource intensity of existing tools. These limitations often result in mission drift, where startups prioritize survival over their original goals. To address this, a new tool has been developed that simplifies impact measurement for startups, allowing them to report their contributions to societal change transparently and efficiently. Thus, the principal implication of this study is that new firms should engage in ESG activities because doing so will not just improve their legitimacy, but it may also help them obtain more funding from investors.

Keywords Impact measurement · Sustainable entrepreneurship · Design science · Investor assessment

JEL Classification L26 · L31 · Q56 · E22 · G11

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11187-025-01101-7.

C. Eckerle (✉) · O. Terzidis
Karlsruhe Institute of Technology, Karlsruhe, Germany
e-mail: christin.eckerle@kit.edu

O. Terzidis
e-mail: orestis.terzidis@kit.edu

Published online: 04 August 2025

1 Introduction

Measuring startups' social and environmental impact is central to assessing the extent of the economic and social impacts that these companies can have in promoting positive change. Schäfer et al. (2015) define sustainable entrepreneurship as an umbrella term encompassing all impact-driven startups that aim to facilitate transformational change by generating social or ecological value (Eckerle et al., 2024). However, there is a debate on whether impact-driven startups should aim to measure outcomes and impact, as the causal link between outputs and outcomes is often unclear and beyond the company's control (Crucke & Decramer, 2016). Ebrahim and Rangan (2014) argue that measuring impacts and outcomes may be counterproductive for a single company, suggesting that these should be measured at an aggregated level across multiple companies or at the country level. This is supported by research suggesting that the utilization of environmental, social, and governance (ESG) criteria (output level) is useful, as evidence shows that companies with prior ESG experience and high ESG ratings are more likely to comply with engagement requests, leading to higher impact at later stages (Kölbel et al., 2020).

Reporting on sustainability at the output level might help startups to be more successful in acquiring investments. The study by Truong and Nagy (2021) suggests that potential investors may be more inclined to invest in new venture opportunities when they perceive sustainability initiatives as a source of long-term competitive advantages for the nascent ventures they are considering. While research highlights that measuring sustainability increases their legitimacy as an investment case (Lall, 2019; Scherer et al., 2013), startups constantly struggle with resource restrictions that constrain them to activities essential to their organizations' emergence and viability (Truong & Nagy, 2021).

This highlights the challenges faced by impact-driven startups with hybrid goals (Battilana, 2018; Santos et al., 2015). Impact-driven entrepreneurs frequently encounter goal conflicts and seek ways to manage them (Pieniazek et al., 2024; Vedula et al., 2022). However, if impact measurement is perceived as too complex or resource intensive, mission drift might be the outcome of these goal conflicts, and there may be less and less alignment between business activities and impact (Eckerle & Terzidis, 2024; Santos et al., 2015).

Current impact measurement tools exhibit high complexity. Many of these tools are not tailored to the startup context (Horne, 2019). For example, they require data as evidence, posing a challenge for startups with limited capacity and resources (Fichter et al., 2023; Singhania & Swami, 2024). Additionally, the chosen impact indicators must be robust enough to prove the startup's impactfulness (Lall, 2019).

This study aims to synthesize the existing body of knowledge to provide design support for new instrumental artifacts to foster startups' impact assessments, addressing the aforementioned constraints of startups. The following research question is addressed:

How to design new instrumental artifacts to provide support for the impact assessment of impact-driven startups?

To answer this question, a design science research (DSR) approach is followed (Romme & Dimov, 2021; Romme & Reymen, 2018; Terzidis et al., 2023). Based on derived design requirements (DRs) and design principles (DPs), an impact assessment tool has been developed through three design cycles for impact-driven startups, aiming to measure and report their potential impact in addressing societal problems and facilitate a transparent assessment by impact investors.

The proposed solution advocates adopting common measurement models and indicators to foster transparency and comparability. The instrumental artifact incorporates essential startup needs, facilitating the measurement and reporting of impact in a resource-adequate manner while accounting for the overarching goal of impact investors to generate positive societal impact. Empowering startups to provide a transparent report of their potential impact aims to increase their legitimacy. The instrumental artifact is designed to streamline decision-making for impact investors and other stakeholders, accounting for entrepreneurial capacities.

The remainder of this paper is structured as follows. First, a short overview of existing impact measurement and assessment approaches is given. Second, the design science methodology is introduced, providing an overview of the approach as well as the theoretical foundations for the design iterations. Then, the main results are presented, followed by an in-depth presentation of the three evaluation cycles.

Third, the paper critically discusses the findings and concludes with a summary, an overview of the limitations, and suggestions for future research.

2 Theoretical background

Many startups aim for economic goals, such as rapid and exponential growth, as well as significant market shares. Such startups, which rapidly increase the company's value and attract the interest of investors and venture capitalists, are referred to as unicorns (Kollmann et al., 2020).

In contrast to this startup type, impact-driven startups prioritize ecological and social objectives in addition to economic ones (Schäfer et al., 2015). Raising capital is a much greater challenge for such startups, especially in direct comparison to their growth-driven peers (Hazenberg et al., 2013). Therefore, impact investing is an attractive funding source for these types of startups, as impact investors emphasize the impact goal alongside a (reduced) financial return (Hockerts et al., 2022). However, impact investors demand measurable evidence of the potential impact the startup intends to achieve (Bocken, 2015; Grieco et al., 2015).

Multiple ongoing efforts are being made to develop and standardize measurement tools for impact-driven companies. Prominent organizations leading this endeavor include the formerly known as "Impact Management Project," now Impact Frontiers (ImpactFrontiers, 2024) and the Global Impact Investing Network (GIIN), which has established the "IRIS+" metric (GIIN, 2019).

Research has identified a demand in practice for developing sound measurement and reporting tools to assess the non-financial sustainable performance of companies (Crucke & Decramer, 2016). Having originated in social science and aid programs (Roth, 2021), several impact measurement tools have been developed in recent years that focus on the application in a business context (Rainock et al., 2018). However, there is a lack of suitable measurement tools for startups, as they have distinct needs and may face different challenges than established companies (Fichter et al., 2023; Horne, 2019).

Impact measurement approaches can be subdivided into three groups: process-based, scorecard-based, and synthetic indicators (Arena et al., 2016; Horne, 2019).

First, the process-based approach aims to establish a causal relationship between the startup's actions and the resulting change or impact. Evidence in the academic literature shows that process-based approaches are suitable for qualitatively reporting the intended impact (Roth, 2021). Prominent examples adjacent to the "theory of change" (Grieco, 2015; Wendt, 2021) include the "social return on investment" (Nicholls et al., 2012) and the "social balanced scorecard" (Grieco et al., 2015).

Second, the scorecard approach aims to evaluate the sustainable impact of the startup across various dimensions. The sustainability performance or potential of the startup in the different dimensions is rated on a predefined scale (e.g., 1–5 points). The scores for each dimension are then added up, multiplied, and combined to produce a final score. This final score enables comparison with other startups, previous scores, or the highest score. Examples of such dimensions include environment, value to society, scalability, and efficacy (Malhotra et al., n.d.). They offer a straightforward interpretation that is beneficial in accountability processes (O'Flynn & Barnett, 2017). Two well-known and established scorecard-based models are the "impact compass" (Malhotra et al., n.d.) and the "B impact assessment" by BLab (Barman, 2015).

Third, synthetic indicators compare the impact of different organizations (Grabenwarter & Liechtenstein, 2011; Horne, 2019). While the scorecard approach aims to assess the sustainability impact of the startup internally, synthetic indicators aim to compare the impact across different organizations (Horne, 2019). Therefore, it is an approach suitable for investors. Here, a single indicator is calculated that combines several aspects of a company's impact. Such indicators aim to integrate the various aspects of impact into a comprehensive measure that is easier to communicate and understand. Therefore, KPIs are often used that can be adapted to the needs and characteristics of a particular impact investment (Grabenwarter & Liechtenstein, 2011). Here, one noteworthy framework to compare different companies is the "global impact investing rating system" (GIIRS). GIIRS is an IRIS+-based rating approach for investors. It is used to rate companies on different dimensions (e.g., employees, operations, and environment) using the IRIS + metrics (Barman, 2015). The aim is to make it easier to measure impact and ensure the results are comparable. However, this approach has also been criticized for being too superficial in terms of metrics and therefore does not allow for an adequate impact assessment (Brest & Born, 2013).

In summary, all the artifacts used in these three approaches (process, scorecard, or synthetic indicators) have different problems that are very versatile and multilateral. They include high resource requirements, a lack of standardization and normalization, and poor comparability (Arena et al., 2016; Horne, 2019). This highlights the entry barrier to impact measurement, especially from a startup perspective, which in turn hinders their assessment by impact investors.

3 Methodology

A DSR approach is employed to investigate design knowledge and develop design impact assessment tools that facilitate the assessment and comparability of impact-driven startups. This research approach has been proven in entrepreneurship research to be fitting in addressing complex challenges and bridging the gap between theory and practice (Terzidis et al., 2023). The approach by Kuechler and Vaishnavi (2008) presents a structured method enabling the generation of design knowledge. Notably, it strongly emphasizes practical implementation and effective communication of the artifact. In this respect, the recommendations by Kuechler and Vaishnavi (2008) are utilized to iteratively develop an instrumental artifact that allows testing, refining, and evaluating the developed design requirements and derived design principles from the studies mentioned above in this paper.

This study's research is structured according to the phases of awareness of the problem, suggestion, development, evaluation, and conclusion in three design cycles (see Fig. 1).

In the initial phase, designated as the "awareness of problem" stage, the problem area is delineated, with a foundation in both extant literature and practical considerations. For instance, emergent trends within the industry or insights derived from an associated discipline may serve as the initial point of inquiry. The subsequent phase, "suggestion," directly follows the first and develops guidelines for a preliminary design solution. In this creative process, the problem is considered from different angles, allowing certain aspects to be set aside and others to be emphasized. This step confirms whether the proposal from the first phase has merit or whether it should be discarded. The subsequent phase, "development," involves implementing the tentative design, contingent upon the creation of the artifact. This phase is followed by an "evaluation," where rigorous testing is ensured, providing empirical evidence of the artifact's utility and efficacy. The final phase, entitled "conclusion," involves the synthesis of key findings, the acknowledgment of limitations, and the proposal of future steps (Kuechler & Vaishnavi, 2008).

The starting point of this research is the derivation of DRs and related DPs, resulting from a research synthesis of previous studies (Denyer et al., 2008). The formulation and validation of DPs that encompass the generic capabilities of designed artifacts help to build a cumulative body of design knowledge (Baskerville & Pries-Heje, 2010).

Following the framework by Gregor et al. (2020), with a specified context (C), aim (A), mechanism (M), and rationale (R), this logic is applied to present the defined DP adequately, ensuring structured design support for future design projects concerning the impact assessment of startups. Thereby, the context specifies the application of an artifact in a

Design Science Cycle	Design Cycle 1	Design Cycle 2	Design Cycle 3
Awareness	Synthesis of DRs	Findings from DC 1	Findings from DC 2
Suggestion	Derived DPs	Refinement of DPs	Refinement of DPs
Development	Instantiation of artifact	Instantiation of improved artifact	Instantiation of improved artifact
Evaluation	Interviews with experts	Interviews with users	Expert feedback, focus group
Conclusion	Interview analysis	Interview analysis	Evaluation analysis, final artifact

Fig. 1 Overview of the iterative design cycles (own illustration)

specific setting. The aim addresses the target group(s) and highlights the intended effect of the artifact. The mechanism then describes a concrete action to achieve the intended effect. Lastly, the rationale justifies the necessity of the mechanism based on theoretical knowledge. This logic has been applied previously in developing design principles for, for example, sustainability assessment in business model innovation (Bhatnagar et al., 2022).

A key evaluation proposed by Venable et al. (2012) is determining the artifact's utility for the intended purpose. While evaluation studies typically test the artifact when a known means-end relationship already exists, they can also evaluate more abstract design principles (Seckler et al., 2021). Therefore, the instrumental artifact was evaluated three times at the end of the design cycles to validate the derived DPs for impact assessment of startups. The three design cycles, along with their respective evaluation strategies, are presented briefly.

The *first design cycle* aimed to translate current knowledge into a tangible prototype for assessing the impact of early-stage startups. The defined design requirements and the derived set of theory-grounded design principles (cf. Section 4.1) guided the process, describing the generic functions and capabilities of the artifact. In particular, the design principles build on social learning theory with its inherent tendency of homophily (Gangopadhyay & Nilakantan, 2021; Sun & Tang, 2011), the principal-agent theory with the problem of information asymmetry (Eckerle & Terzidis, 2024; Reid, 1999), and the conflicting goal theory of impact-driven entrepreneurs (Pieniazek et al., 2024; Vedula et al., 2022).

The design principles are instantiated in a first instrumental artifact, a paper prototype (V0), which showcases the generic composition and central elements (see Fig. 5, cf. Section 4.6). This prototype version was evaluated via 20 semi-structured interviews with investors, consultants, startups, and established companies experienced in impact investing or sustainable innovation and financing (see Supplementary Material, Experts).

The evaluation focused on the overall setup of the instrumental artifact, including its main elements, and the proposed impact criteria that were included in the artifact. The interviews were semi-structured and lasted, on average, 45 min. They began with questions about the experts' understanding and experience with impact measurement, followed by feedback on the different elements and criteria of the prototype. The expert interviews were recorded, transcribed, and deductively analyzed, according to the building blocks of the artifact, as well as inductively coded, following the recommendations of Mayring (2014). The results confirmed the effectiveness of a process- and scorecard-based model and revealed opportunities for further improvements of the instrumental artifact, particularly in the impact categories.

The *second design cycle* aimed to enable the overall usability of the artifact from the startups' perspective. Based on the results from the first cycle and in line with former research on impact measurement and ESG criteria, the overall goal shifted to a more output-level perspective on reporting indicators. Hence, the impact criteria were adapted to better fit the possibilities of startups in measuring their potential impact. The prototype was then transferred to the first version of the tool (V1). It was built on an Excel sheet with the tool's description and a template the startup could fill out. In this way, creating a situation similar to reality allowed for testing the artifact's usability. The supplementary material, including the data logbook, showcases the evolution of the different versions, as V1 only differs slightly from the final version (V3), see Fig. 3 for the main setup (cf. Section 4.2). The evaluation in this cycle was conducted through in-depth interviews with three startups that had already been interviewed in the previous cycle.

One startup is active in the healthcare sector, one in climate tech, and one supports companies that want to integrate sustainability processes into their existing business activities. The startup's data is anonymized for confidentiality reasons. The startups worked with the impact measurement template; researchers recorded their verbal feedback. Again, the recorded interviews were transcribed and deductively coded, following the recommendations of Mayring (2014). The results confirmed the validity of the proposed instrumental artifact. They only suggested minor changes to the artifact, such as changing the CO2 equivalent indicator to be measured in tons per Euro instead of kilograms.

The *third design cycle* focused on assessing and comparing multiple startups by investors (V2). Prompted by the results of the previous startup user testing, the final qualitative feedback on the usefulness and applicability of the tool was retrieved from impact investors, impact accelerators, and impact investing consultants. For the feedback, an Excel sheet was prepared, containing 1) a worksheet with an overview of the evaluation goal, a step-by-step instruction for using the evaluation sheet, and the assessment tool; 2) a description of all the elements of the startup measurement tool; 3) three fictive startup cases showcasing a filled-out impact measurement template; 4) a benchmarking of the three fictive startup cases for comparison; 5) An evaluation worksheet for feedback. Again, this approach was chosen as it is suitable, as it simulates a real situation, enabling applicants to imagine using the tool in daily situations, evaluating impact-driven startups, and comparing them with other impact-driven startups. The investor perspective feedback was obtained from four impact investors, two impact investing consultants, and one impact accelerator. The acquisition was based on the initial expert panel from cycle one, followed by a snowball screening of other impact investors (Goodman, 1961). The written feedback was deductively coded, following Mayring's (2014) guidelines. Overall, the experts confirmed the applicability of the instrumental artifact.

Lastly, a focus group workshop (Morgan, 1997) was conducted with six researchers in the field of sustainability and impact investing at scientific conferences to confirm the scientific robustness of the artifact. They were briefly introduced to all the elements, including the measurement aspects of the startups and the benchmarking aspects for investors. By collectively reflecting on the artifact, new thoughts were inspired and triggered (Szopinski et al., 2019). On a virtual whiteboard, a researcher collected feedback concerning strengths, weaknesses, and points for improvement, while a second researcher moderated the open discussion. These categories also informed the subsequent coding of the data. Only minor suggestions lead to the refinement of the tool, resulting in its final version (V3). As the results of the third evaluation cycle did not prompt any significant changes to the content of the design solution, this concluded the process.

The overall results of the iterative design process, along with the main findings of the evaluations, are described below.

4 Results

The first part of this section summarizes and describes the various requirements for an impact assessment tool for startups, as well as the design principles. The final instrumental artifact, along with its main components, is presented in the second part. The third part introduces the results of the evaluations.

4.1 Design requirements and proposed design principles

The overarching goal of the impact assessment tool for impact-driven startups is to minimize the liability of impactfulness, thereby making it a more legitimate investment case for impact investors (Lall, 2019).

In three design cycles, based on insights from theory, prior research, and three evaluation episodes, four design principles are proposed to inform the design of impact assessment tools that support impact investors in evaluating a legitimate startup case. The design principles are derived from seven design requirements, informed by previous research, as shown in Table 1. Their reasoning will be elaborated in the following.

A study by Eckerle et al. (2024) focused on the underlying criteria for identifying an impact-driven startup. The findings offer several implications for the design support for impact assessment tools. First and foremost, entrepreneurs should articulate their impact mission in the context of a business case, aligning it with the Sustainable Development Goals (SDGs) and providing comprehensive research on the severity of the societal problem they are addressing. Using commonly known models, such as a logic model (e.g., a theory of change), allows stakeholders to evaluate the startup's claims more easily and enhances legitimacy (DR1, DR2). Clarifying the activities to achieve desired outcomes will likely improve monetary and nonmonetary valuation by impacting investors and other supporting organizations (Harrer & Owen, 2022). Since former research indicates that criteria are difficult to quantify or predict (Crucke & Decramer, 2016), entrepreneurs should strive to present their impact as concretely as possible. This can minimize information asymmetry between investors and startups (cf. Eckerle & Terzidis, 2024).

As Eckerle et al. (2024) showed, it was possible to identify a consolidated and validated set of

Table 1 Design requirements for a startup impact assessment tool (own illustration)

Number	Description
DR1	The artifact should integrate a clear and precise logic model showcasing how the startup intends to impact a societal problem positively
DR2	The artifact should incorporate widely accepted standards within the impact investing community regarding the models and the indicators included
DR3	The artifact should include the comprehensive concept of sustainability while allowing for case-specific criteria
DR4	The artifact should require a small number of impact categories to be measured by the startup to reduce complexity
DR5	The artifact should ensure a comprehensive process for impact measurement that can be communicated lucidly
DR6	The artifact should enable the comparison of impact-driven startups between predefined categories

assessment criteria applicable to impact investing. The results suggest considering only a few criteria when assessing legitimate investment cases. Nevertheless, all synthesized impact criteria are essential for sustainable development and should not be neglected. A positive impact and the avoidance of negative externalities should be ensured (Schutselaars et al., 2023), advocating a holistic measurement of sustainability activities. This may minimize the likelihood of impact washing (Azmat et al., 2022). However, investors may lean towards well-known and common criteria (Kollmann & Kuckertz, 2010), which can be explained by social learning theory, suggesting a tendency towards homophily (Eckerle et al., 2024). To make a holistic contribution to the SDGs, investors need to consider all consequences, including social, environmental, and economic (Elkington, 1997) (DR3), which can help minimize the risk of mission drift among investors (Cetindamar & Ozkazanc-Pan, 2017).

Startups face specific challenges and barriers in measuring and reporting their potential impact to external stakeholders (Fichter et al., 2023; Siefkes et al., 2023; Singhania & Swami, 2024). Research highlights the need to simplify the complexity of the impact measurement process while ensuring clear communication of the impact (Horne, 2019) (DR5). Further, setting a clear goal by utilizing the SDGs for target setting captures the startup's unique goal, allowing for case specification (Schutselaars et al., 2023) (DR3). At the same time, clearly defined and a low number of impact indicators on the company level can enable certain comparability within and between startups (Roundy et al., 2017) (DR6), ensure holistic sustainability (Siefkes et al., 2023) (DR3), and reduce resource constraints (Hirschmann et al., 2022; Pierrakis & Owen, 2023; Singhania & Swami, 2024) (DR4).

The six design requirements inform four design principles as design knowledge for developing an impact assessment tool. Further, three underlying theories inform their rationale: the conflicting goals of impact-driven entrepreneurs (Pieniazek et al., 2024; Vedula et al., 2022), social learning with the underlying tendency of homophily (Gangopadhyay & Nilakantan, 2021; Sun & Tang, 2011), and principal-agent theory with the problem of information asymmetry (Eckerle & Terzidis, 2024; Reid, 1999). Over the three design cycles, the design principles were constantly refined. The final set of the four design principles is presented in Table 2.

All four design principles have the same context: they present design knowledge for impact assessment tools in the context of *impact investing in startups*.

DP1, DP2, and DP3 aim to minimize the liability of impactfulness and enhance the legitimacy of impact-driven startups, while DP4 aims to enable an impact investor to compare potential investment cases. The first three design principles address the aforementioned problem of conflicting goals (Pieniazek et al., 2024), which could lead to a startup's mission drift. This informs the mechanism of DP1, which proposes to *integrate measurable indicators based on state-of-the-art standards to align a startup's business activities with its potential impact*. DP2 adds to this by proposing a mechanism for *determining impact indicators that represent the comprehensive concept of sustainability, while considering startup-specific capabilities to measure their potential impact*. The latter part highlights the conflicting goals of impact-driven startups. In contrast, the first part of this DP acknowledges the tendency towards

Table 2 Design principles for startup impact assessment tools (own illustration)

Number	Aim (A)	Context (C)	Mechanism (M)	Rationale (R)
DP1	To minimize the liability of impactfulness and enhance the legitimacy of an impact-driven startup	for impact investment in startups	integrate measurable indicators based on state-of-the-art standards to align business activities of a startup with its potential impact	as it helps to manage conflicting goals within the impact-driven startup
DP2	To minimize the liability of impactfulness and enhance the legitimacy of startups	for impact investment in startups	determine impact indicators representing the comprehensive concept of sustainability while considering startup-specific capabilities to measure their potential impact	because this reduces the tendency of social learning based on homophily by impact investors while accounting for conflicting goals of impact-driven entrepreneurs
DP3	To minimize the liability of impactfulness and enhance the legitimacy of startups	for impact investment in startups	ensure an easily understandable and slim information synthesis of the impact claim for both investor and startup	because this accounts for conflicting goals of impact-driven entrepreneurs and reduces information asymmetry
DP4	To enable an impact investor to compare potential investment cases	for impact investment in startups	ensure the consistent use of indicators by the investor for their impact investing decisions	because this reduces the information asymmetry

homophily, which ultimately informs social learning (Murnieks et al., 2011), as evident in the investment behavior of impact investors (Eckerle et al., 2024).

Lastly, the third principle proclaims a mechanism for *ensuring an easily understandable and slim information synthesis of the impact claim for investors and startups*. Again, this accounts for the conflicting goals of impact-driven startups and reduces information asymmetry between the investor and the startup, a crucial aspect of principal-agent theory that is evident in investor-investee relationships (Reid, 1999). This reduction in information asymmetry is substantiated by the mechanism of DP4, which suggests *ensuring the consistent use of indicators by investors for their impact investing decisions.*

Figure 2 illustrates the connection between the developed design requirements and the final set of design principles, which are represented by their mechanisms, their relationship to the target groups, and the underlying theories.

Given these requirements and principles for the design solution, along with the three evaluation cycles, the final design solution (V3), outlined in the next section, was developed. Overarching is the startup's intended contribution to the SDGs. The impact assessment tool incorporates a scorecard and a process model to bolster this claim. Only a restricted number of impact criteria are required, and the measurement indicators are focused on the output level. However, they encompass all three dimensions of sustainability—social, environmental, and ecological—to ensure a comprehensive assessment of the startups.

4.2 Final instantiation

The instrumental artifact consists of two components: the startup's measurement and reporting side and the investor's assessment and comparison side. For the startup to fill out the tool, the final design solution (V3) begins with a scorecard-based model, as the data can later be utilized to understand claims made in the logic model (Jackson, 2013; Nachyla & Justo, 2024), as shown in Fig. 3. In addition to the filled-out tool, the investor will receive a cumulated table of all potential investment cases, with a benchmarking chart to visually present the differences between the cases (Fig. 4). The following section will provide a detailed introduction to the main components of the artifact, as well as the key findings of the evaluations.

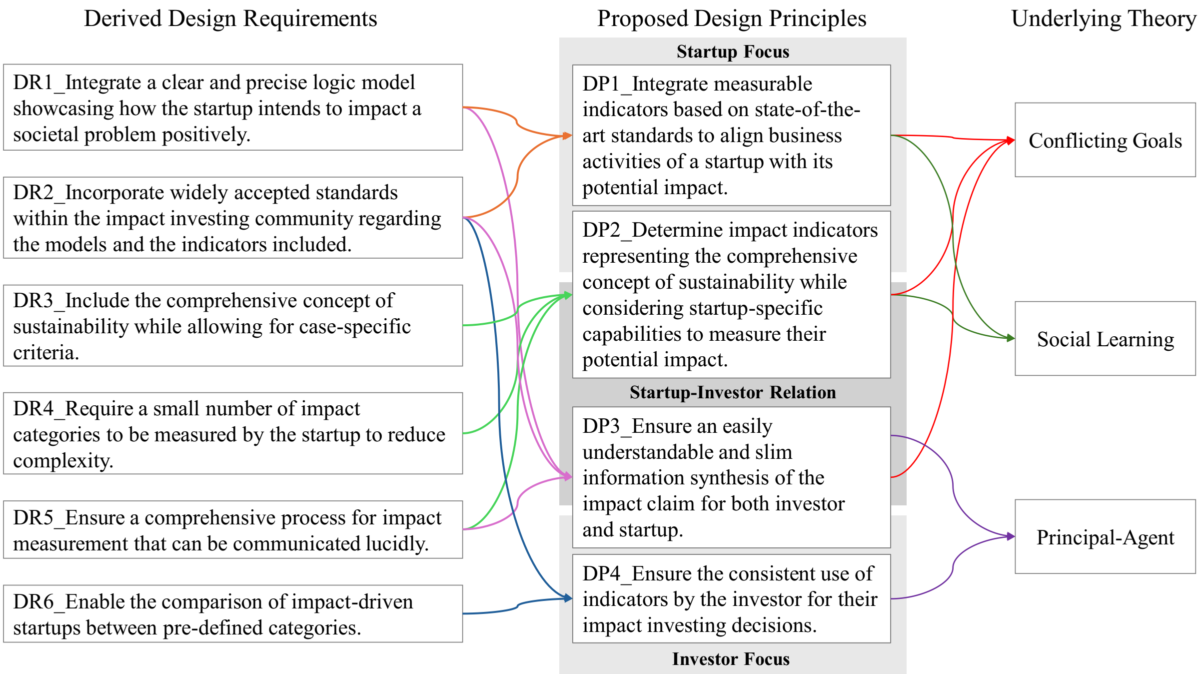

Fig. 2 Derived DRs and final DPs with their underlying theories (own illustration)

Input

Revenue from current year		in €
Revenue from year before		in €
Scope 1 / Scope 2 Greenhouse Gas Emissions in CO2-equiv. (current year)		in tons
Scope 1 / Scope 2 Greenhouse Gas Emissions in CO2-equiv. (year before)		in tons

CAGR Revenue Development	**Values**	
Number of Periods		
Initial Value		in €
Final Value		in €

Rate of Impact Investments	**Values**	
Amount of Impact Investments		in €
Amount of Invested Capital		in €

Diversity (in %)	**Values**
Number of people within the organization who feel that they belong to minority groups	
Number of women within the organization	
Number of employees	

Affected SDGs & related target(s)	**Description**
Description of how the activities of the Startup contribute to the targeted SDG(s). Provide a logic model of your intended process.	* see space below output table

Output

Dimension	**Score**
Scope 1 Greenhouse Gas Emissions in CO_2-equiv.	0
Revenue in Relation to Scope 1 Greenhouse Gas Emissions in CO_2-equiv.	0
Difference Greenhouse Gas Emissions in CO2-equiv. between current year and the year before	0
CAGR Revenue Development (in %)	=(FV/IV)^(1/Nr.ofPeriods)-1
Rate of Impact Investments (in %)	0,00%
Diversity I: Percentage of people within the organization who feel that they belong to minority groups (in %)	0,00%
Diversity II: Percentage of women within the company (in %)	0,00%

Logic Model	Affected SDG(s)
For (who or what is the beneficiary) **who/which currently face** (state the problem) **we intend to** (state potential positive impact) **by** (how do you plan to do this)	SDG xy; Target yz
…	…

Fig. 3 Artifact version 3—startup template with exemplary case (own illustration)

4.3 Measuring scorecard-based indicators

The startup-related reporting starts with the measurement of scorecard-based indicators. To account for startups' resource constraints, three criteria reflecting the triple bottom line of sustainability are proposed to be measured via eleven indicators. Overall, these indicators are similar to standard ESG criteria, which have been proven to contribute to the achievement of the SDGs (Betti et al., 2018) and account for the overall sustainability efforts of the young company. This enhances legitimacy, as these criteria are commonly recognized and further outline potential weaknesses if a startup is less active or produces negative externalities in one of the dimensions (cf. Schutselaars et al., 2023).

For the scorecard-based measurement, the startup's revenue from the current year serves as the basis for subsequent impact-related indicators. For the potential environmental impact, the startup reports on its scope 1 (or scope 2, if data is available) CO2 emissions (measured in tons) from the previous year to the current year.

Economic sustainability is typically reported via the compound annual growth rate (CAGR), a method that visualizes a company's average annual growth rate over a specified period. CAGR can be applied to a wide range of economic indicators and is, therefore, characterized by a high degree of flexibility (Gartner, n.d.):

$$CAGR = \left(\frac{Final\ value}{Starting\ value}\right)^{\frac{1}{N}} - 1$$

The number of years to be analyzed, n, is determined by the year change. The quotient of the final value and the initial value describes the total growth over the period under review. This is distributed over the years using the compound interest effect (nth root). Investors have a vested interest in the long-term trajectory of economic development. In this context, the following parameters are often considered: turnover, profit, loss reduction, and customer base development. Following the investment, the investor anticipates a growth trajectory for the

Output				**Output - normalized Data**		
Dimension	**Score Case 1**	**Score Case 2**	**Score Case 3**	**Case 1**	**Case 2**	**Case 3**
Scope 1 / Scope 2 Greenhouse Gas Emissions in CO_2-equiv., in tons	200,0	100,0	6000,0	2,3	2,0	3,8
Scope 1 / Scope 2 Greenhouse Gas Emissions in CO2-equiv. in relation to revenue	5000,0	490,0	1333,3	3,7	2,7	3,1
Positive Difference of Greenhouse Gas Emissions in CO2-equiv. between current year and year before, in tons	30,0	10,0	1700,0	1,5	1,0	3,2
CAGR Revenue Development (in %)	58,1	121,4	100,0	1,8	2,1	2,0
Rate of Impact Investments (in %)	75,0	16,7	66,7	1,9	1,2	1,8
Diversity I: Percentage of people within the organization who feel that they belong to minority groups (in %)	23,1	20,0	42,0	1,4	1,3	1,6
Diversity II: Percentage of women within the company (in %)	30,8	100,0	54,0	1,5	2,0	1,7
SDGs Adressed	13	12	10; 12	13	12	10; 12

Fig. 4 Artifact version 3—investor benchmarking with exemplatory cases (own illustration)

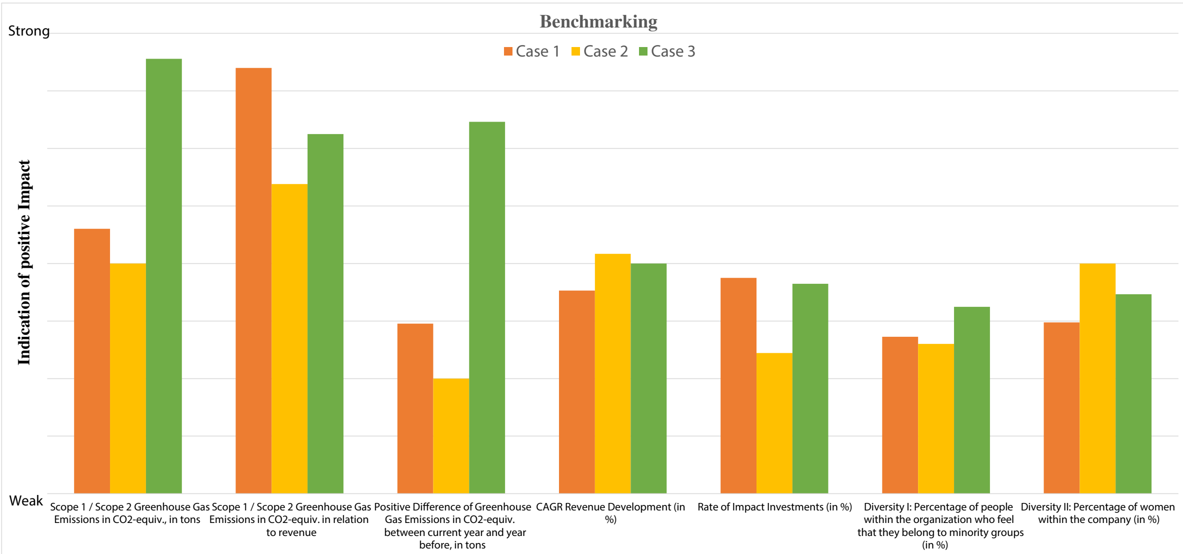

Fig. 4 (continued)

startup over the coming years. The prerequisite for this calculation is market entry or initial sales profits. However, this is a relatively straightforward calculation for nascent businesses, as they typically generate sales before achieving profitability.

To connect the business growth potential to the scalability of impact, the proportion of the company's value invested in generating impact represents the startup's management of conflicting goals. Impact investments refer to both the amount of investment already received and the company's value if a startup has not yet received any investment. The aim is to obtain a percentage figure that indicates the amount of money flowing into projects or activities that generate impact. Projects or activities can also be the product itself. Generally, these activities are aligned with the SDGs (see Section 4.4). In the calculation, the amount of money spent is divided by the amount of money received. This allows the investor to compare how much of the invested capital is ultimately invested in impact.

To measure a startup's social sustainability, the diversity of its employees is evaluated. Three indicators are used to measure diversity: the total number of employees and the percentage of women in a company are measured, as gender equality is essential, for example, for securing EU funding or attracting female-oriented venture capitalists (Yang et al., 2020). The second indicator focuses on the "percentage of people within the organization who feel that they belong to minority groups." Minority groups in this regard are, for example, socio-economic background, immigration, language barrier, disability, sexual orientation, or gender (other than female/male). Here, anonymized surveys can be used to measure this.

4.4 Reporting the theory of change

While the former output-level indicators represent the internal view of a startup's activities towards a potential positive impact, the tool incorporates an external view via its connection to the SDGs, utilizing a logic model that explains how the startup's activities are intended to align with the respective SDG. Previous research has highlighted that the SDGs are a widely accepted and preferred framework for understanding a startup's impact claims (Burckart et al., 2018; Schutselaars et al., 2023). At the same time, the description of the startup's process, including its activities and intended outcomes or impact, acknowledges the specific requirements of each startup case.

The respective column in the tool is called "Affected SDG(s) and related target(s)," which describes how the startup's activities contribute to its targeted SDG(s). To account for resource constraints, for easy communication, and to still achieve certain comparability between startup cases, the logic model is divided into the following parts: "For (who or what is the beneficiary) who/which currently face (state the problem), we intend to (state potential positive impact) by (how do you plan to do this)."

This comprises the theory of change, which is based on the most important facts: the beneficiary and the state of the problem demonstrate a direct relation to the addressed SDG. At the same time, the potential positive impact and the plan for achieving it represent the impact or outcomes, along with their related inputs and activities. The structure of this condensed logic model builds on earlier entrepreneurship research concerning the value proposition for impact measurement (Schutselaars et al., 2023), which is a one-sentence formulation emphasizing the importance of conciseness in translating the core value of a startup (Manthey, 2024). The value proposition statement is commonly used in practice, making it easy to use and understand for both impact-driven startups and impact investors.

4.5 Benchmarking

Lastly, the tool provides investors with a table and a chart visualization to help them quickly and easily understand how the different startups perform concerning their potential for positive impact, which becomes visual and thereby interpretable due to the depiction of multiple startup cases (see Fig. 4). On the y-axis, the potential for positive impact ranges from weak to strong. On the x-axis, the different dimensions from the startup measurement tool are presented and listed based on the environmental, economic, and social groups. Therefore, the visual representation ensures the investor's attention to a holistic picture concerning the potential impact, which counters the tendency of homophily. The tool's aim is not to represent the different sectors of the startups nor to represent the stage of development of the startups, recognizing the limitations of this comparison. Still, to a certain extent, comparison is possible by normalizing the available data. Overall, the benchmarking chart helps assess each startup's potential to have a positive impact in specific categories compared to other investment cases.

4.6 Evaluation of the tool

The tool and its earlier versions were tested and validated using various evaluation methods. The tool's initial prototype (V0) was reviewed via 21 expert interviews. This evaluation led to the refinement and development of the first applicable tool (V1). For the user testing of the V1, four startups were invited to apply the tool, and researchers were present to record the startups' opinions and suggestions. Based on their feedback, V2 was developed, incorporating a benchmarking tool that enables investors to compare different startup cases. The third evaluation cycle consisted of 1) qualitative expert feedback and 2) a focus group workshop with researchers in management, sustainable finance, and entrepreneurship. Based on this, the final version (V3) has been developed. An overview of the main changes over each iteration, along with exemplary quotes, is presented in Table 3. More detailed data on the different evaluation cycles can be found in the Supplementary Material (data logbook). All data is anonymized due to confidentiality reasons.

The results of the first evaluation episode revealed that the initial prototype of the instrumental artifact (see Fig. 5), which combined a process and scorecard model, was perceived as reasonable. The developed design requirements (cf. Section 4.1) for measuring startup impact laid the groundwork for designing prototype V0.

Therefore, the prototype elements consist of the most relevant SDG a startup aims to address, matched with the impact goal(s), followed by a theory of change to showcase the activities of the startup aiming towards the impact goal(s). Secondly, to ensure comparability between startups while accounting for individual startup specifics, a web-cobb depicts the scorecard measurement of impact categories with set indicators that the startup has to measure. Here, based on the findings from the survey with impact investors (Eckerle et al., 2024), the prototype includes impact criteria for the environmental (saved CO_2-equivalent, biodiversity via renatured land), social (diversity, employee engagement), and economic (return on sales, number of employees) dimension of sustainability, as well as one user-definable criterion which is aligned with the main impact goal of the startup. If possible, the criteria are aligned with the startup's revenue to ensure a match between business growth and impact potential.

Table 3 Overview of evaluation results and decisions for changes (own illustration)

Design cycle	Version	Data source	Goal	Main feedback	Main decisions for next version
DC1	V0	Expert interviews with investors, consultants, startups, and established companies	Translate current knowledge into a tangible prototype for impact assessment of early-stage startups	Logic model in combination with the SDGs is fitting to justify and explain the startup's overall impact goal Impact indicators demand changes due to high complexity, missing relevance in the startup context, specificity Economic indicators should reflect rate of impact investments	Input and evaluation table implemented to ensure connection between values and to highlight main output Highlighting connection of business activities to targeted SDG(s) Scorecard chart excluded due to exclusion of comparison with ideal state Multiple changes of indicators
DC2	V1	User testing with startups	Enable the overall usability of the artifact from the startups' perspective	The rate of impact investments received favorable reception. Still, threats or weaknesses stem from the lack of comparability of this criterion, and the question of concrete impact activities vs. necessary business activities Minor adaptions needed to other indicators in scorecard Shift focus to startups that have already generated revenue for at least one year	Sales changed to revenue, inclusion of comparison between two years of scope 1 greenhouse gas emissions in CO2-equivalent in tons instead of kg Adapted logic model based on well-known value statement in accordance with the theory of change model
DC3	V2	Expert feedback from investors and intermediaries Focus group workshop with researchers	Ensure the assessment and comparability of multiple startups by impact investors	Trade-off between clarity and complexity More and more users in different sectors can increase comparability ESG-level indicators only provide an indication of potential impact Measuring only turnover/profit in the economic area may not be sufficient Suggestion of using some choice of essential vs. necessary conditions to choose from Need for reconsideration of target group/startup age	Revenue between two years can be compared directly in the tool Startups can choose between scope 1 or scope 2, based on data availability Adapted visualization of benchmark chart to enhance clarity for comparison

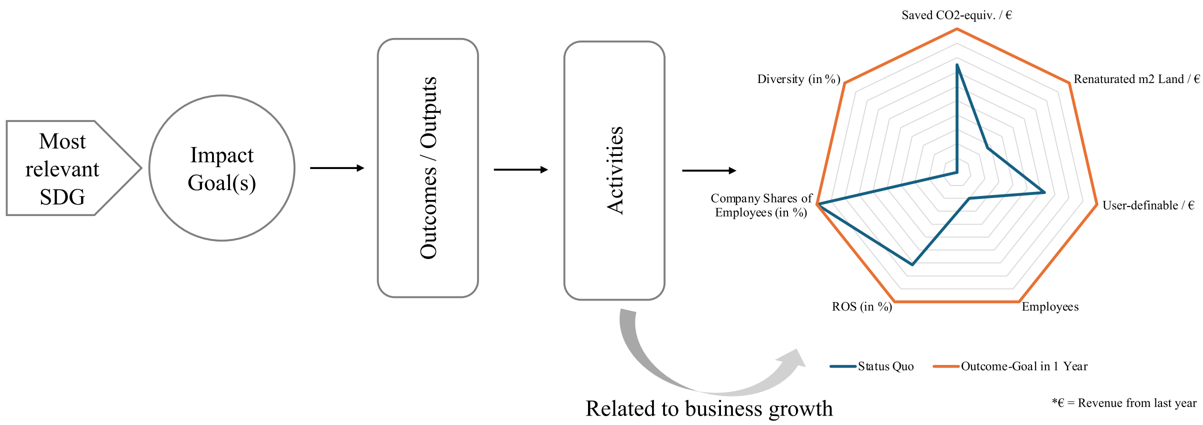

Fig. 5 Prototype of the startup impact assessment tool (own illustration)

While the experts deemed a logic model, such as the theory of change, in combination with the SDGs, to be fitting for justifying and explaining the startup's overall impact goal, major points for improvement emerged for the scorecard-based indicators. This proved to be the red threat in all evaluation cycles: which indicators are relevant, essential, or measurable for startups.

In the first evaluation round, the experts proposed reducing the number of impact criteria from 5 to 3. It is particularly noteworthy that the "user-definable" impact criterion has been discarded, since this has the effect of hindering comparability. According to the experts, allowing startups too much freedom in defining key performance indicators (KPIs) might lead to irrelevant or less meaningful metrics, which could undermine the integrity and comparability of the assessment. The same fate befell the "biodiversity" criterion, as it was disregarded as too complex to measure and too case-specific. Here, experts highlighted that obtaining accurate data for biodiversity-related metrics is challenging, as companies often lack precise information on land usage and restoration. Further, biodiversity as a KPI may be too niche and not universally relevant, particularly for sectors with limited direct environmental impact, such as software startups. Lastly, the criterion "return on sales" was rejected, as the experts questioned its necessity in analyzing the potential impact. In the economic dimension, it was proposed that the money spent by the startup on direct impact would be more interesting. This led to the "rate of impact investment" criterion. The aim is to obtain a percentage figure that indicates how much money flows into projects or activities by the startup that generates impact. Only minor suggestions for improvements were made for the other impact criteria; overall, experts accepted them.

The suggestions from this evaluation led to the development of the first applicable tool (V1). The focus of this iteration was on the usability for startups. The user testing of V1 confirmed that the overall setup is easy to understand. Furthermore, the indicators for the three impact criteria were discussed thoroughly and further refined. No essential changes were made regarding the replacement, addition, or renaming of the criteria; however, various weaknesses were identified to facilitate the startup application. The insights the startups provided helped refine the criteria to better align with their capabilities, necessitating primarily adjustments to the units. For example, for the CAGR criterion and the match between CO2 equivalent and saved tons per Euro, startups need to have generated a certain revenue to make sense. The startups themselves suggested that making a speculative forecast could negatively impact investors' ability to compare startup cases. Therefore, the focus of the assessment tool should be on startups that have already generated revenue for at least one year.

Overall, the connection to the SDG was highlighted as a significant benefit of acknowledging startups' individual goals. To account for a comprehensible representation of the logic model, a refinement of the in-depth presentation of the theory of change was proposed. A deep dive into the entrepreneurship education literature guides this refinement (Manthey, 2024), see Section 4.4. Concluding the second evaluation phase, the tool was deemed appropriate for measuring and assessing the startups' impact at a lifecycle stage of at least one year of revenue.

The last refinements were incorporated into the second version of the tool (V2). Based on the finalized output table from the startup measurement tool, a benchmark representation was designed, as the third design cycle focused on assessing and comparing multiple startup cases by impact investors. This version was evaluated with two target groups: impact investors and consultants, as well as accelerators and researchers.

The qualitative feedback from impact investors, consultants, and accelerators suggests that the tool is a good starting point for assessing and comparing the potential impact of startups. A main critique focused on the output level of the data measured by the startup. It became evident that the experts perceive the tool as only partially allowing one to assess the impact of a startup; rather, it showcases the potential for impact. One suggestion to include outcome-level data was to include scope 2 CO2 emission measurement for startups. However, some startups might be unable to retrieve the data for scope 2, as they depend on third-party data (Santos & Eisenhardt, 2009). Therefore, the final version offers this as an optional indicator.

Another point for questioning was raised regarding the diversity measure, as some experts perceived it as very vague, and each company might interpret and report it differently (if at all). However, no concrete solution for improvement was given, resulting in no changes to this dimension.

Notably highlighted was that the tool provides a concise and helpful overview of the potential positive impact of a startup, encompassing all dimensions of sustainability. It was acknowledged that the tool is easy to use and understand. The clarity of the tool was highlighted, and it was pointed out that it strikes a good balance between informative value and time expenditure. Furthermore, the benchmarking approach was positively perceived by some experts, as it offers substantial value, especially when applied across portfolios or funds. However, the internal benchmarking lacks an objective comparison of impact, as the data is provided at the output level.

The results of the focus group workshop with the six researchers highlight that the tool simplifies a complex reality, which is a strength and a weakness simultaneously. Due to the expertise of the experts, it fosters opportunities for startups and investors to learn and consider the impact. At the same time, the output level of impact measurement limits the informative value. Building on these remarks, the results can be clustered into strengths, weaknesses, and suggestions for further improvements.

Notably, the holistic consideration of impact was positively seen, as some startups might only be aware of specific impacts. Furthermore, the small number of indicators was considered a positive aspect, which can be seen as a starting point for impact measurement. These make it easy to use and simplify the complex endeavor of impact measurement. As mentioned, a weakness is the lack of consideration for the impact a startup might generate outside the company, which would be the outcome or even the impact level of measurement, for example, scope 1 CO2 emissions versus scope 2, which represents the outcome level of activities by a startup. Overall, however, the researchers acknowledged the difficulty in balancing impact measurement and startup capabilities, which led to the development of an output-based design for the tool. To strengthen comparability for impact investors, it was suggested that the assessment tool be used only between similar types of startups, such as those with a specific impact focus, active in the same sector, or at a similar stage of development. Furthermore, it may be challenging for young startups to report on these indicators, suggesting that the tool may only be applicable after two or three years of operation. For this, the researchers suggested qualitative reporting on the indicators to support the numbers with written explanations, which can be combined with the theory of change.

In conclusion, the last evaluation indicates that the tool is complete and useful for measuring the output level of sustainability activities in a startup, indicating the potential for impact. Moreover, the tool appears to support impact investors to a certain extent in assessing and comparing multiple startups. The combination of indicators and qualitative reporting on activities proved valuable, highlighting how the startup aims to contribute to the SDGs. In terms of the requirements defined in Section 4.1, it can thus be concluded that the tool fulfills all requirements formulated, which in turn supports the derived DPs.

5 Discussion

The design instantiation offers a potential solution to an instrumental artifact for assessing the impact of startups on impact investors. The design solution has been meticulously constructed, drawing on a multitude of scholarly works and interviews, with the aim of comprehensive incorporation without compromising quality. What must be highlighted about the evaluations is the persistent debate about the measurement of output level versus outcome or impact level—an ongoing debate in theory and practice (Singhania & Swami, 2024). This highlights the struggle to acknowledge the resource constraints of startups while demanding sound and transparent evidence of impact.

Nevertheless, new initiatives in practice confirm the proposed solution by choosing similar approaches. One example is the initiative “Score-4Impact,” a collaborative project involving research institutions, public impact initiatives, and investment companies, now a gGmbH (BMWK, n.d.); score-4impact, n.d.). They propose the following approach, which aligns closely with the design instantiation. Their initial step for startups involves an ESG measurement focusing on five environmental, five social, and five governance criteria, supplemented with educational components and implementation aids. This includes scoring based on the current implementation status and improvement potential, categorized by sector, resulting in a strengths-weaknesses profile of the ESG measures and evaluation of basic disclosure principles. The second step, the “impact estimator,” defines impact goals, estimates greenhouse gas reduction potentials, aligns with SDGs, and compares the solution’s emission savings to established alternatives (Score4Impact, 2024).

This showcases once more the feasibility of ESG-related measurement, which can strengthen the legitimacy of startups’ impact claims. This is likewise evident in the alignment with the SDGs. Furthermore, the reduced number of indicators acknowledges the adjustment of the impact assessment tool to startup resource constraints. The design solution further incorporates the connection to the startup’s revenue. Therefore, the goal of impact investors to align a startup’s business activities with its intention to generate a positive impact is taken into account. This interconnection and direct relation between impact and revenue can minimize the risk of mission drift and help manage conflicting goals (Pieniazek et al., 2024; Santos et al., 2015).

The connection to revenue illustrates the tool’s restriction: it is only applicable for startups with at least one year of revenue. It is suggested that less-mature startups begin by reporting on the logic model. If revenue is provided, both aspects—the output-level measurement connected to the revenue and the alignment with the SDGs—ensure certain comparability while acknowledging individual cases. The comparison can be increased if two startups address the same SDG or operate in similar sectors. Lastly, the chosen indicators and design ensure attention to potential weak aspects, providing a holistic sustainability picture (Kaufmann & Botha, 2024).

6 Conclusion

The existence and relevance of assessment tools for impact are commonly acknowledged. Therefore, our study focused on identifying overarching principles to design rigorous and relevant artifacts that provide support for the impact assessment of impact-driven startups. An optimal solution was iteratively designed to test and validate the DPs applicable to startups while ensuring comparability for impact investors. The DSR approach, as proposed by Kuechler and Vaishnavi (2008), served as a guideline for the scientific advancement of the design instantiation.

Specific to design science, the contribution of this study is mainly theoretical. Based on the framework developed by Seckler et al. (2021), this research has contributed to the theoretical design of object knowledge in impact assessment for startups. Seckler et al. (2021) define theoretical design object contribution as, for example, design science studies at a high level of abstraction that contributed to the development of DPs. This study derived and iteratively validated four DPs, presented once more here in a comprehensive form.

DP1 integrates measurable indicators based on state-of-the-art standards to align the business activities of a startup with its potential impact. *DP2* determines impact indicators representing the comprehensive concept of sustainability while considering startup-specific capabilities to measure their potential impact. *DP3* ensures an easily understandable and slim information synthesis of the impact claim for both investor and startup. *DP4* ensures the consistent use of indicators by the investor for their impact-investing decisions.

The DPs make a significant contribution to design knowledge that can guide the future development of such instrumental artifacts for the impact assessment of startups. They cover all necessary perspectives: the startup capabilities (DP1), the startup-investor relation (DP2, 3), and the investor side (DP4).

The evaluation focus informs parts of the study's limitations. First, only a few user tests were conducted for the application of the impact measurement tool by startups. Therefore, testing the impact measurement tool with more startups would be beneficial. Hereby, the selection of startups should pay attention to include startups of different sizes, at various life cycle stages, and from diverse industries to evaluate whether the tool can be applied to all kinds of startups or is more suitable for specific industries.

Second, the instrumental artifact was only evaluated in artificial settings (Venable et al., 2016). This was partly due to the difficulties in acquiring suitable interview and user testing partners. Impact investors, in particular, proved once more to be a challenging target group to reach, which is why other stakeholders, such as consultants or accelerator program managers, were included in the development of the final artifact. However, not all stakeholders have a say or are involved in selecting possible ventures for investments. As a result, the final artifact should be again evaluated by impact investors. This aligns with suggestions from DSR for follow-up research, which involves empirically testing the proposed design solution for its effectiveness (Seckler et al., 2021; Venable et al., 2016).

Author contribution Both authors contributed to the study conception and design. Material preparation and data collection were performed by Christin Eckerle, and both authors performed the analysis. The first draft of the manuscript was written by Christin Eckerle; both authors commented on previous versions of the manuscript. Both authors read and approved the final manuscript.

Funding Open Access funding enabled and organized by Projekt DEAL.

Data availability If not given in the supplementary material, the data are available per the authors' request.

Declarations

Ethical approval Informed written consent has been obtained from all interview partners involved in the study.

Conflict of interest The authors declare no competing interests.

References

Arena, M., Azzone, G., Bengo, I., & Calderini, M. (2016). Indicators and metrics for social business: A review of current approaches. *Journal of Social Entrepreneurship, 7*(1), 1–24. https://doi.org/10.1080/19420676.2015.1049286

Azmat, F., Jain, A., & Michaux, F. (2022). Strengthening impact integrity in investment decision-making for sustainable development. *Sustainability Accounting, Management and Policy Journal, 13*(1), 55–87. https://doi.org/10.1108/SAMPJ-10-2020-0368

Barman, E. (2015). Of principle and principal: Value plurality in the market of impact Investing. *Valuation Studies, 3*(1), 9–44. https://doi.org/10.3384/vs.2001-5592.15319

Baskerville, R., & Pries-Heje, J. (2010). Explanatory design theory. *Business & Information Systems Engineering, 2*(5), 271–282. https://doi.org/10.1007/s12599-010-0118-4

Battilana, J. (2018). Cracking the organizational challenge of pursuing joint social and financial goals: Social enterprise as a laboratory to understand hybrid organizing. *M@n@gement*, *21*(4), 1278–1305.

Betti, G., Consolandi, C., & Eccles, G. R. (2018). The relationship between investor materiality and the Sustainable Development Goals: A methodological framework. *Sustainability, 10*(7), 2248. https://doi.org/10.3390/su10072248

Bhatnagar, R., Keskin, D., Kirkels, A., Romme, A. G. L., & Huijben, J. C. C. M. (2022). Design principles for sustainability assessments in the business model innovation process. *Journal of Cleaner Production, 377*, 134313. https://doi.org/10.1016/j.jclepro.2022.134313

BMWK. (n.d.). Score4Impact - Klimaschutz bei Start-ups investierbar machen. In *Nationale Klimaschutzinitiative*: Bundesministerium für Wirtschaft und Klimaschutz.

Bocken, N. M. P. (2015). Sustainable venture capital - catalyst for sustainable startup success? *Journal of Cleaner Production, 108*, 647–658. https://doi.org/10.1016/j.jclepro.2015.05.079

Brest, P., & Born, K. (2013). *When can impact investing create real impact?* www.ssireview.org

Burckart, W., Lydenberg, S., & Ziegler, J. (2018). *Measuring effectiveness: Roadmap to assessing system-level and SDG investment*. https://sdg-investments.com/fileadmin/user_upload/Measuring-Effectiveness-Report_Final_March-2018.pdf

Cetindamar, D., & Ozkazanc-Pan, B. (2017). Assessing mission drift at venture capital impact investors. *Business Ethics*, *26*(6). https://doi.org/10.1111/beer.12149

Crucke, S., & Decramer, A. (2016). The development of a measurement instrument for the organizational performance of social enterprises. *Sustainability*, *8*(2), 161–161. https://doi.org/10.3390/SU8020161

Denyer, D., Tranfield, D., & Van Aken, J. E. (2008). Developing design propositions through research synthesis. *Organization Studies, 29*(3), 393–413. https://doi.org/10.1177/0170840607088020

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. https://doi.org/10.1525/cmr.2014.56.3.118

Eckerle, C., & Terzidis, O. (2024). Designing impact due diligence for startups. *Journal of Business Venturing Design, 3*, 100020–100020. https://doi.org/10.1016/j.jbvd.2024.100020

Eckerle, C., Finner, A. S., & Terzidis, O. (2024). Investor perspective on impact-driven startups–the prioritization of certainty. *Social Enterprise Journal, 20*(4), 611–631. https://doi.org/10.1108/SEJ-02-2024-0028

Elkington, J. (1997). The triple bottom line. *Environmental management: Readings and cases*, *2*.

Fichter, K., Lüdeke-Freund, F., Schaltegger, S., & Schillebeeckx, S. J. D. (2023). Sustainability impact assessment of new ventures: An emerging field of research. *Journal of Cleaner Production, 384*, 135452–135452. https://doi.org/10.1016/j.jclepro.2022.135452

Gangopadhyay, P., & Nilakantan, R. (2021). Peer effects and social learning in banks' investments in information technology. *International Review of Economics & Finance, 75*, 456–463. https://doi.org/10.1016/j.iref.2021.03.024

Gartner (n.d.). Compound annual growth rate. In *Gartner glossary*.

GIIN. (2019). *IRIS+ and the five dimensions of impact*.

Goodman, L. A. (1961). Snowball Sampling. *The Annals of Mathematical Statistics, 32*(1), 148–170. https://doi.org/10.1214/aoms/1177705148

Grabenwarter, U., & Liechtenstein, H. (2011). In search of gamma an unconventional perspective on impact investing. In *IESE Business School Working Paper*.

Gregor, S., Chandra Kruse, L., & Seidel, S. (2020). Research perspectives: The anatomy of a design principle. *Journal of the Association for Information Systems*, *21*(6), 1622–1652. https://doi.org/10.17705/1jais.00649

Grieco, C. (2015). *Assessing social impact of social enterprises. Does one size really fit all?* SpringerBriefs in Business. http://www.springer.com/series/8860

Grieco, C., Michelini, L., & Iasevoli, G. (2015). Measuring value creation in social enterprises: A cluster analysis of social impact assessment models. *Nonprofit and Voluntary Sector Quarterly, 44*(6), 1173–1193. https://doi.org/10.1177/0899764014555986/FORMAT/EPUB

Harrer, T., & Owen, R. (2022). Reducing early-stage Cleantech funding gaps: An exploration of the role of Environmental Performance Indicators. *International Journal of Entrepreneurial Behavior & Research, 28*(9), 268–288. https://doi.org/10.1108/IJEBR-10-2021-0849

Hazenberg, R., Seddon, F. A., & Denny, S. (2013). Intermediary perceptions of investment readiness in the social investment market. *Seventeenth Annual Conference of the International Research Society for Public Management (IRSPM XVII)*, 10–12. http://nectar.northampton.ac.uk/5181/

Hirschmann, M., Moritz, A., & Block, J. H. (2022). Motives, supporting activities, and selection criteria of social impact incubators: An experimental conjoint study. *Nonprofit and Voluntary Sector Quarterly, 51*(5), 1095–1133. https://doi.org/10.1177/08997640211057402

Hockerts, K., Hehenberger, L., Schaltegger, S., & Farber, V. (2022). Defining and conceptualizing impact investing: Attractive nuisance or catalyst? *Journal of Business Ethics, 179*, 937–950. https://doi.org/10.1007/s10551-022-05157-3

Horne, J. (2019). The sustainability impact of new ventures. Measuring and managing entrepreneurial contributions to sustainable development. In.

ImpactFrontiers. (2024). Impact frontiers: Impact for investment decision-making. In.

Jackson, E. T. (2013). Interrogating the theory of change: Evaluating impact investing where it matters most. *Journal of Sustainable Finance and Investment, 3*(2), 95–110. https://doi.org/10.1080/20430795.2013.776257

Kaufmann, L., & Botha, H. (2024). Who loses in win-win investing? A mixed methods study of impact risk. *Journal of Business Ethics*. https://doi.org/10.1007/s10551-024-05788-8

Kölbel, J. F., Heeb, F., Paetzold, F., & Busch, T. (2020). Can sustainable investing save the world? Reviewing the mechanisms of investor impact. *Organization & Environment, 33*(4), 554–574. https://doi.org/10.1177/1086026620919202

Kollmann, T., Jung, P. B., Kleine-Stegemann, L., Ataee, J., & de Cruppe, K. (2020). *Deutscher Startup Monitor 2020. Innovation statt Krise* (9783948895044).

Kollmann, T., & Kuckertz, A. (2010). Evaluation uncertainty of venture capitalists' investment criteria. *Journal of Business Research, 63*(7), 741–747. https://doi.org/10.1016/j.jbusres.2009.06.004

Kuechler, W., & Vaishnavi, V. (2008). On theory development in design science research: Anatomy of a research project. *Proceedings of the Third International Conference on Design Science Research in Information Systems and Technology*. https://www.researchgate.net/publication/220393130

Lall, S. A. (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. https://doi.org/10.1007/sll266-018-00081-5
Malhotra, N., Clavier, B., & Bailard, E. (n.d.). *The impact compass*. https://www.gsb.stanford.edu/experience/about/centers-institutes/csi/impact-compass
Manthey, S. I. (2024). Adopting design science research to develop a framework for systematic technology application selection. In. Karlsruhe.
Mayring, P. (2014). Qualitative content analysis: Theoretical background and procedures. In *Approaches to qualitative research in mathematics education: Examples of methodology and methods* (pp. 365–380). Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-017-9181-6_13
Morgan, D. (1997). *Focus Groups as Qualitative Research*. SAGE Publications, Inc. https://doi.org/10.4135/9781412984287
Murnieks, C. Y., Haynie, J. M., Wiltbank, R. E., & Harting, T. (2011). 'I like how you think': Similarity as an interaction bias in the investor-entrepreneur dyad. *Journal of Management Studies, 48*(7), 1533–1561. https://doi.org/10.1111/j.1467-6486.2010.00992.x
Nachyla, P., & Justo, R. (2024). How do impact investors leverage non-financial strategies to create value? An impact-oriented value framework. *Journal of Business Venturing Insights*, *21*. https://doi.org/10.13039/501100011033
Nicholls, J., Lawlor, E., Neitzert, E., & Goodspeed, T. (2012). *A guide to social return on investment*. The SROI Network. https://socialvalueselfassessmenttool.org/wp-content/uploads/intranet/758/pdf-guide.pdf
O'Flynn, P., & Barnett, C. (2017). *Evaluation and impact investing: A review of methodologies to assess social impact*.
Pieniazek, R., Unsworth, K. L., & Dean, H. (2024). How and why do social entrepreneurs experience goal conflict differently? *Journal of Business Venturing Insights, 21*, 452–452. https://doi.org/10.1016/j.jbvi.2024.e00452
Pierrakis, Y., & Owen, R. (2023). Startup ventures and equity finance: How do Business Accelerators and Business Angels' assess the human capital of socio-environmental mission led entrepreneurs? *Innovation, 25*(4), 371–395. https://doi.org/10.1080/14479338.2022.2029706
Rainock, M., Everett, D., Pack, A., Dahlin, E. C., & Mattson, C. A. (2018). The social impacts of products: A review. *Impact Assessment and Project Appraisal, 36*(3), 230–241. https://doi.org/10.1080/14615517.2018.1445176
Reid, G. C. (1999). The application of principal-agent methods to investor-investee relations in the UK venture capital industry. *Venture Capital, 1*(4), 285–302. https://doi.org/10.1080/136910699295820
Romme, A. G. L., & Dimov, D. (2021). Mixing oil with water: Framing and theorizing in management research informed by design science. *Designs, 5*(1), 13. https://doi.org/10.3390/DESIGNS5010013
Romme, A. G. L., & Reymen, I. M. M. J. (2018). Entrepreneurship at the interface of design and science: Toward an inclusive framework. *Journal of Business Venturing Insights*, *10*(e00094). https://doi.org/10.1016/j.jbvi.2018.e00094
Roth, K. (2021). An analysis on impact measurement: How do we measure impact? In K. Wendt (Ed.), *Theories of change: Change leadership tools, models and applications for investing in sustainable development* (pp. 139–152). Springer International Publishing. https://doi.org/10.1007/978-3-030-52275-9_10
Roundy, P., Holzhauer, H., & Dai, Y. (2017). Finance or philanthropy? Exploring the motivations and criteria of impact investors. *Social Responsibility Journal, 13*(3), 491–512. https://doi.org/10.1108/SRJ-08-2016-0135
Santos, F., Pache, A.-C., & Birkholz, C. (2015). Making hybrids work: Aligning business models and organizational design for social enterprises. *California Management Review, 57*(3), 36–58. https://doi.org/10.1525/cmr.2015.57.3.36
Santos, F. M., & Eisenhardt, K. M. (2009). Constructing markets and shaping boundaries: Entrepreneurial power in nascent fields. *Academy of Management Journal, 52*(4), 643–671.
Schäfer, K., Corner, P. D., & Kearins, K. (2015). Social, environmental and sustainable entrepreneurship research: What is needed for sustainability-as-flourishing? *Organization and Environment, 28*(4), 394–413. https://doi.org/10.1177/1086026615621111
Scherer, A. G., Palazzo, G., & Seidl, D. (2013). Managing legitimacy in complex and heterogeneous environments: Sustainable development in a globalized world. *Journal of Management Studies, 50*(2), 259–284. https://doi.org/10.1111/joms.12014
Schutselaars, J., Romme, A. G. L., Bell, J., Bobelyn, A. S. A., & Van Scheijndel, R. (2023). Designing and testing a tool that connects the value proposition of deep-tech ventures to SDGs. *Designs 2023*, *7*(2), 50–50. https://doi.org/10.3390/DESIGNS7020050
Score4Impact. (2024). *Discover the next ESG & climate champion. Tools & benchmarking für Investoren, Förderer & Startups*.
score4impact. (n.d.). *score4impact - unser Ansatz*. https://score4impact.com/unser-ansatz/
Seckler, C., Mauer, R., & Vom Brocke, J. (2021). Design science in entrepreneurship: conceptual foundations and guiding principles. *Journal of Business Venturing Design*, *1*(1–2). https://www.journals.elsevier.com/journal-of-business
Siefkes, M., Bjørgum, Ø., & Sørheim, R. (2023). Business angels investing in green ventures: How do they add value to their startups? *Venture Capital*, 1–30. https://doi.org/10.1080/13691066.2023.2260101
Singhania, M., & Swami, D. (2024). Impact investing: Scientometric review and research agenda. *Business Ethics, the Environment and Responsibility, 33*(3), 251–286. https://doi.org/10.1111/beer.12599
Sun, J., spsampsps Tang, J. (2011). A survey of models and algorithms for social influence analysis. In *Social Network Data Analytics* (pp. 177–214). Springer. https://doi.org/10.1007/978-1-4419-8462-3_7
Szopinski, D., Schoormann, T., & Kundisch, D. (2019). Theoretical foundation and empirical evidence view project

visual languages for business model representation View project Cover sheet because your taxonomy is worth it: Towards a framework for taxonomy evaluation. https://www.researchgate.net/publication/332711034

Terzidis, O., Gutmann, M., & Ziegler, C. (2023). Design science research in entrepreneurship: Scaffolding artifacts for the venturing process. https://doi.org/10.5445/IR/1000162284

Truong, Y., & Nagy, B. G. (2021). Nascent ventures' green initiatives and angel investor judgments of legitimacy and funding. *Small Business Econom, 57*(4), 1801–1818. https://doi.org/10.1007/s11187-020-00373-5

Vedula, S., Doblinger, C., Pacheco, D., York, J. G., Bacq, S., Russo, M. V., & Dean, T. J. (2022). Entrepreneurship for the public good: A review, critique, and path forward for social and environmental entrepreneurship research. *Acad Manag Ann, 16*(1), 391–425. https://doi.org/10.5465/annals.2019.0143

Venable, J., Pries-Heje, J., spsampsps Baskerville, R. A (2012). A comprehensive framework for evaluation in design science research. In K. Peffers M. Rothenberger, Kuechler B (Eds.), *Design science research in information systems. Advances in theory and practice. DESRIST 2012. Lecture Notes in Computer Science* (Vol. 7286, pp. 423–438). Springer. https://doi.org/10.1007/978-3-642-29863-9_31

Venable, J., Pries-Heje, J., & Baskerville, R. (2016). FEDS: A framework for evaluation in design science research. In *European Journal of Information Systems*. (Vol. 25, pp. 77–89): Palgrave Macmillan Ltd.

Wendt, K. (2021). Theory of change: Defining the research agenda. In K. Wendt (Ed.), *Theories of change: Change leadership tools, models and applications for investing in sustainable development* (pp. 3–12). Springer International Publishing. https://doi.org/10.1007/978-3-030-52275-9_1

Yang, S., Kher, R., & Newbert, S. L. (2020). What signals matter for social startups? It depends: The influence of gender role congruity on social impact accelerator selection decisions. *Journal of Business Venturing*, *35*(2). https://doi.org/10.1016/j.jbusvent.2019.03.001