

Article

The Development of a Measurement Instrument for the Organizational Performance of Social Enterprises

Saskia Crucke * and Adelién Decramer

Ghent University, Faculty of Economics and Business Administration, Tweeckerkenstraat 2, 9000 Ghent, Belgium; Adelién.Decramer@UGent.be

* Correspondence: Saskia.Crucke@UGent.be; Tel.: +32-9-2432956; Fax: +32-9-2432909

Academic Editor: Giuseppe Ioppolo

Received: 16 December 2015; Accepted: 2 February 2016; Published: 6 February 2016

Abstract: There is a growing consensus that the adoption of performance measurement tools are of particular interest for social enterprises in order to support internal decision-making and to answer the demands of accountability toward their stakeholders. As a result, different methodologies to assess the non-financial performance of social enterprises are developed by academics and practitioners. Many of these methodologies are on the one hand discussions of general guidelines or, on the other hand, very case specific. As such, these methodologies do not offer a functional tool for a broad range of social enterprises. The goal of this article is to fill this gap by developing an instrument suitable for the internal assessment and the external reporting of the non-financial performance of a diverse group of social enterprises. To reach this goal, we used qualitative (focus groups and a Delphi panel) and quantitative research methods (exploratory and confirmatory factor analysis), involving multiple actors in the field of social entrepreneurship. Focusing on five dimensions of organizational performance (economic, environmental, community, human and governance performance), we offer a set of indicators and an assessment tool for social enterprises.

Keywords: performance measurement; social enterprises; social performance

1. Introduction

Due to the growing interest in sustainability and the organizational responsibilities to society, organizations face the challenge of assessing and reporting their non-financial performance. This is especially the case for social enterprises [1–3]. Social enterprises are social mission-driven organizations that develop an entrepreneurial activity (make products and/or deliver goods and services) in order to fulfill unsolved social needs in society [4,5]. They are considered as a distinct category of organizations, positioned between profit and nonprofit organizations [6,7]. Social enterprises differ from profit organizations as profit is not a goal as such, but a mean to create social value [8]. Compared to nonprofit organizations, social enterprises establish entrepreneurial activities to ensure their financial sustainability and rely not (exclusively) on subsidies and donations [9].

Because of the dual mission of creating social value and being financially sustainable, financial as well as non-financial performance are core to the social enterprise functioning [10]. Social enterprises are described as typical hybrid organizations [11–13] and face some specific internal and external tensions and challenges [10]. This challenging environment has made the assessment and the reporting of the organizational performance within social enterprises of particular importance. Firstly, different authors warn of internal tensions because of the difficulty of balancing the financial and social goals in decision-making, and refer to mission drift, which is the erosion of the social goals in favor of financial goals, as a threat [9,14,15]. In addition to the annual account—which is useful to evaluate the financial performance—a tool that supports social enterprises to assess and discuss internally

their non-financial performance might be helpful in balancing the social and the financial goals in decision-making and in avoiding mission drift [2,8]. Secondly, social enterprises face external tensions, related to the need to establish legitimacy and to obtain support from different stakeholder groups [16,17]. As social enterprises lack a dominant external stakeholder, they are exposed to multiple and conflicting expectations and demands of different principal stakeholder groups [14]. The legitimacy perceived by stakeholders is crucial for resource acquisition, such as financial resources and human resources [9,13]. Important stakeholder groups of social enterprises are the beneficiaries of the social mission and the customers, paying for the products and services delivered by the social enterprise [14]. Further, also policy makers, funders and volunteers can have a legitimate stake in the organization. These stakeholders expect assessments and reporting to be transparent and comparable [18,19]. This has brought social enterprises under significant pressure to seek ways to more actively manage and report their non-financial performance to answer the demands of accountability to multiple stakeholders [17]. As such, social enterprises feel a need to develop, implement and use a systematic approach to assess and report their non-financial organizational performance [8].

Although there is a consensus that the development and adoption of performance measurement instruments is of particular interest for social enterprises, there is a lack of methodologies with a practical usefulness for a broad range of social enterprises [3,17]. This paper aims at describing the development of a set of indicators and an assessment tool, useful for evaluating and reporting the non-financial organizational performance of social enterprises. While the financial performance of social enterprises can be evaluated based on the information available in the annual account of the organization, the aim of this paper is to develop a tool to assess the non-financial performance of social enterprises. If we refer, in the following sections, to the organizational performance of social enterprises, we actually focus on the non-financial organizational performance.

The structure of the paper is as follows. In the following section, we briefly review relevant literature on performance measurement in social enterprises. Next, we describe the different steps carried out to identify relevant indicators and to develop an assessment tool to assess organizational performance in social enterprises. The article concludes with a discussion of the development of this assessment tool, challenges involved in its use and suggestions for future research.

2. Performance Measurement in Social Enterprises

The idea that organizations should measure and manage their performance is a key-issue in management literature and is strongly encouraged by international bodies such as the OECD and the World Bank [20]. There is a growing consensus that social enterprises should assess their performance to support internal decision-making and to respond to the increasing demands of accountability towards different stakeholders [3,21]. As a result, researchers and practitioners have developed different methodologies and tools to measure the performance of social enterprises [22]. These methodologies and tools are however diverse and make a comparison of the organizational performance of social enterprises very difficult [8]. Grieco, Michelini and Iasevoli [2] (p.1) state that “the overall picture remains fragmentary if not confusing”. The reason why methodologies and tools are falling short of expectations is twofold. On the one hand, some studies are “general” in their design and do not offer specific indicators or measurement tools. The developed methodologies and tools often discuss frameworks providing general guidelines for social enterprises considering designing and implementing a performance measurement system, e.g. Manetti [1]. These papers discuss for instance how diverging stakeholder expectations can be taken into consideration or they present different steps that social enterprises can follow to implement a performance measurement system [8]. They offer as such no insight in the dimensions or indicators that can or should be evaluated [2]. Other papers discuss relevant dimensions of organizational performance (e.g. environmental performance, social performance, *etc.*), but do not propose relevant performance indicators [23]. On the other hand, other studies are too specific and are examining performance measurement in specific cases, and make it difficult to replicate and generalize to other social enterprises. Bellucci, *et al.* [24], for instance, study the

performance of fair trade shops in Italy. The performance indicators studied are specifically related to the fair trade value chain and cannot be replicated without adaptations to other organizational contexts.

Differences in approach and methods related to performance measurement in social enterprises can be attributed to two antecedents. Firstly, social enterprises differ in size, activities, objectives and accordingly relevant stakeholders. By consequence, it is not easy to develop a model that is suitable to all kind of social enterprises [8]. Secondly, performance measurement can serve different purposes. Generally spoken, performance measurement can have an internal or an external purpose. A performance measurement tool can be used as an internal management instrument, enabling organizations to assess their performance and support internal decision-making. On the other hand, performance management tools with an external purpose are used for external reporting and have the main purpose of accountability to stakeholders. A different purpose implies a different design of the performance measurement system [2]. Notwithstanding this diversity in organizations and performance measurement systems, we also notice a consensus on some aspects. First, there is a consensus that organizational performance is multi-dimensional. Not only is there, as mentioned earlier, the difference between financial and non-financial performance. Also, non-financial performance is multi-dimensional taking into consideration performance having an impact on the local community, the environment, society in general, and people working in the organization [2,8]. Secondly, there is a consensus that performance is not only related to immediate results. Many frameworks use a “results chain” or “logic model” [22] also taking in consideration inputs (i.e. resources used) and activities within the organization [8,25]. These models, stressing the alignment of an organization’s input, throughput and output components, can be used to assure program alignment and to evaluate the results of an organization [26]. Concerning the achieved results, a difference is made between immediate results (outputs) and medium and long-term results, often referred to as outcomes or impacts [2]. Although there is growing interest in “impact measurement”, the terms “outcomes” and “impact” are not consistently used [22]. Ebrahim and Rangan [22] distinguish between outputs (immediate results), outcome (medium- and long-term impact on individuals) and impact (medium- and long-term impact on communities or populations). There is a consensus that organizations should at least measure and report on inputs, activities and outputs [20]. However, Ebrahim and Rangan [22] doubt whether social enterprises should go further and also measure outcomes and impact. Their main argument is that the causal link between outputs and outcomes is not clear and that outcomes and impact often go beyond the control of the social enterprise. Some scholars argue that organizations, or the management of these organizations, could be demotivated, withdraw discretionary effort and sit back and see if they win or lose a performance indicator game that resembles a lottery [20]. For instance, a work integration social enterprise offering a job to disadvantaged people can measure the number of people hired by the organization (output). However, whether this will result in an improved quality of life at the individual level (outcome) or a decrease in poverty at society level (impact) is not straightforward as also other external factors beyond the control of the social enterprise will have an influence on this outcome and impact. Furthermore, Ebrahim and Rangan [22] argue, mainly based on practitioner oriented literature, that focusing on the measurement of impacts and outcomes might be counterproductive, because it puts much demand on (often small) organizations without necessarily resulting in better outcomes. Instead, they suggest that outcomes and impacts be measured at an aggregated level, for instance by governments, foundations or impact investors.

Relying on these former insights, the aim of this paper is to develop a performance measurement tool for social enterprises. More specifically, we want to develop a tool suitable for a broad range of social enterprises. Taking in consideration the internal (assessing organizational performance and supporting decision-making) and external (reporting) purpose of performance measurement, we aim at providing social enterprises with a performance measurement instrument which is based on the reliable, valid, and standardized assessment of organizational performance. In developing this tool we build on the insights that performance is multi-dimensional and that, when evaluating performance, inputs, activities and outputs should be considered. Based on the arguments of Ebrahim

and Rangan [22], we will not focus on outcomes and impacts. Moreover, we are convinced that taking into consideration outcomes and impacts will refrain us from developing a tool suitable for social enterprises with diverse activities.

In what follows, we discuss in detail how we developed the performance measurement tool, using qualitative and quantitative research methods and building on the expertise and points of view of a broad range of practitioners in the field of social entrepreneurship.

3. Methodology

The aim of the paper is twofold. On the one hand, we want to identify relevant indicators for assessing the non-financial performance of social enterprises. This set of indicators can serve the external purpose of performance measurement: in the external reporting to stakeholders, social enterprises can elaborate on their non-financial performance related to these indicators. This is in line with existing standards developed to assess non-financial performance (e.g. Global Reporting Initiative, GRI) [27]. These standards offer a broad range of possible performance indicators, and organizations choose, given their activity, performance indicators they consider as relevant. While different efforts have been made to develop specific sets of relevant performance indicators for different kind of organizations, such as NGOs and public sector organizations, this is not the case for social enterprises [28,29].

On the other hand, based on the selected indicators, we want to develop a measurement instrument that social enterprises can use as a self-assessment tool to evaluate internally their non-financial organizational performance. In line with existing quality assessment frameworks, such as in the EFQM Excellence model of the European Foundation on Quality Management, different members of the organization can complete the measurement instrument, enabling the assessment of the non-financial performance and supporting decision-making [30].

In order to be able to realize these two aims, we followed generally accepted guidelines and phases outlined in scale development literature [31,32]. Table 1 gives an overview of the five phases of the research process, combining deductive and inductive methodologies to generate relevant indicators and items [33].

Table 1. Overview of the different phases of the research process.

Research Process	Objectives	Results
Phase 1: Literature review	<ul style="list-style-type: none"> Identify performance domains Identify performance indicators 	<ul style="list-style-type: none"> 5 performance domains 41 indicators
Phase 2: Focus groups	<ul style="list-style-type: none"> Check relevance and completeness of selected performance domains and indicators 	<ul style="list-style-type: none"> 41 indicators (Phase 1) approved 12 additional indicators 53 indicators
Phase 3: Delphi panel	<ul style="list-style-type: none"> Find consensus regarding relevant indicators 	<ul style="list-style-type: none"> 13 indicators removed 40 indicators accepted
Phase 4: Survey instrument development and administration	<ul style="list-style-type: none"> Develop a survey instrument to assess performance on the retained indicators 	<ul style="list-style-type: none"> Survey distributed to 1018 social enterprises Response rate 24% (244 social enterprises)
Phase 5: Validation of relevant indicators and the assessment tool	<ul style="list-style-type: none"> Item reduction Indicator reduction Validation of measurement instrument 	<ul style="list-style-type: none"> Validated measurement instrument 21 indicators

We performed our research in a Belgian region, namely Flanders. As it is the case in Europe in general, Flemish social enterprises mainly emerged because of the persistence of structural unemployment and the need for more active policies to tackle the increasing exclusion of specific groups. These “work integration social enterprises” offer a job to disadvantaged people, but in addition they focus actively on job training, necessary to make reintegration in the labor market possible, and provide social support to solve personal problems which are often obstacles for employment [11,12,34]. While some of these organizations are specifically set up to hire disadvantaged people, others are organizations or local authorities who hire some disadvantaged workers, next to a majority of regular employees [34]. Next to the work integration social economy, there is a growing interest in the entrepreneurial, innovative approach of social enterprises. In Flanders, these social enterprises often are member-based democratic organizations, mainly adopting the organizational form of “cooperatives” [35,36]. In the phases of the research process, we took into consideration and involved (representatives of) the different groups of social enterprises. This will be discussed more in detail when commenting on Phases 2, 3 and 4 of the research process.

3.1. Phase 1: Literature Review

To obtain an overview of relevant performance indicators, we started with an extensive literature review. This deductive approach is appropriate as measuring the non-financial performance of organizations has gained increasing attention in the literature [33]. While screening scientific journals, we noticed that relying on internationally accepted standards is a common practice when studying non-financial performance of organizations [37]. Digging deeper into these standards, we selected four standards often referred to in the literature and which have proven their worth in practice as well as in scientific research: (1) Kinder, Lydenberg, Domini (KLD) social responsibility rating [37,38], (2) Dow Jones Sustainability Index (DJSI) [39,40], (3) Global Reporting Initiative (GRI) [41,42] and (4) ISO 26000 [43,44]. Table A1 provides an overview of the performance domains considered by each of the four selected standards. Based on this overview, we selected five performance domains that (1) are taken into account by several of these standards and (2) are relevant in the context of social enterprises: economic, human, environmental, community and governance performance. Figure 1 gives an overview of the five selected performance domains.



Figure 1. Overview of the five selected performance domains.

Economic performance is related to the economic conditions supporting a strong financial position, important for the viability of organizations. As such, the focus is not on financial indicators reported

in the annual financial accounts of organizations, but on economic indicators influencing these financial indicators [27,45]. Human performance refers to the relationship of the organization with its workforce [46]. Environmental performance focuses on the efforts organizations make to protect nature [47]. Community performance refers to how organizations deal with their responsibilities in society [48], including the relationships with dominant stakeholders: beneficiaries of the social mission and customers, paying for the delivered products and services. Governance performance refers to “systems and processes concerned with ensuring the overall direction, control and accountability of an organization” [49]. Important issues related to organizational governance are board composition and board behavior [50,51], as well as dealing with stakeholder expectations [52]. Governance performance is a particular performance domain as it is expected that good governance practices have a positive impact on organizational decision-making, in turn positively influencing the other performance domains of the organization [53,54].

In a next step, we detected relevant indicators for measuring the five performance domains. As the performance of social enterprises has only rather recently been examined, the literature and research on the performance of social enterprises is still limited. We therefore decided to screen the literature on research regarding the non-financial performance of organizations in general, which is more extensively studied in the context of Corporate Social Responsibility (CSR), where it is referred to as “social performance” [37]. We started with the examination of 10 high impact management journals (included in the ISI Web of Science), looking for articles with “social” and “performance” in the title in the period 1990–2013. Table A2 gives an overview of the screened management journals. As a result, we found 68 articles. Thirty-three articles were not relevant because they did not focus on the non-financial performance of organizations. Analyzing the remaining 35 articles, we concluded that these articles only provide a limited number of possible indicators. The main reason is that many articles do not refer to relevant indicators because they are conceptual or because they use existing ratings provided by, for instance, financial institutions to assess the non-financial performance of organizations.

In a next step, we screened two additional journals: *Journal of Business Ethics* and *Social Enterprise Journal*. We selected these journals because they have a focus on CSR and social enterprises. Once again, we screened the journals for relevant articles with “social” and “performance” in the title in the period 1990–2013 (As “*Social Enterprise Journal*” only exists from 2005, we screened this journal for the period 2005–2013). As a result of this additional screening we found 60 additional articles, which provided us with more relevant indicators.

Based on the literature review, we retained 41 indicators.

3.2. Phase 2: Focus Groups

We noticed that, based on the examined management literature, it is difficult to conclude that the retained indicators are most relevant for the context of social enterprises. Firstly, because the screened literature is not exclusively related to the performance of social enterprises. Secondly, because it is possible that some relevant indicators are not detected in the literature, we therefore decided to combine the deductive approach of the literature review with additional inductive approaches [33]. To check the relevance and completeness of the selected indicators, we organized two focus group sessions. The use of focus groups is a common qualitative research method in social sciences, often as a part of the development of measurement instruments [55]. We decided to organize focus groups because it enables in-depth discussions with experts on emerging and unexplored topics [56].

The reason why we selected key informants with a different background to participate in the focus groups, is twofold. Firstly, we wanted insights on the performance of different kinds of social enterprises (the dominant form of work integration, as well as other social enterprises such as cooperatives). Secondly, we considered it useful to ask the opinion of employees involved in the management of a social enterprise and informants with a broader focus. The latter group are mainly researchers and civil servants supporting social enterprises. We therefore decided to organize two focus

groups. In the first focus group session, eight managers of social enterprises were involved. Because of the prevalence of work integration social enterprises in Flanders [34], we invited managers of different types of work integration social enterprises. In the second focus group, we aim at a broader perspective: three representatives of sectorial federations of work integration social enterprises (The federation of social workshops SST and the federation of sheltered workshops VLAB), two researchers with a broad perspective on social entrepreneurship and two civil servants of the Flemish government were involved.

Organizing two focus groups involving 15 key informants with a different background gave us the opportunity to gain insight into different perspectives regarding measuring the performance of social enterprises. Specifically, we asked the participants whether the five performance domains and the 41 indicators selected in Phase 1 are suitable for assessing the performance of social enterprises and whether there were indicators missing. As a result of the focus groups, the 41 indicators selected in Phase 1 were approved and 12 indicators were added, resulting in 53 indicators. The indicators in each performance domain are presented in Table 2. A distinction is made between indicators selected based on the literature review (Phase 1) and approved by the focus groups (Phase 2), on the one hand, and indicators that are provided by the focus groups (Phase 2), on the other hand.

3.3. Phase 3: Delphi Panel

In focus groups, group dynamics and more particularly the dominance of some participants may substantially influence the results. Moreover, focus groups are not anonymous, potentially making people less outspoken [55,56]. To overcome these potential disadvantages of focus groups, we used the Delphi technique to reach a consensus on the indicators. The Delphi technique encompasses a structured, iterative process in which subject matter experts share their anonymous opinion during subsequent rounds [57–63]. Specifically, this Delphi panel includes 17 panelists with different backgrounds: (1) managers of social enterprises, (2) experts on social entrepreneurship (academics, government officials, representatives of sectorial federations) and (3) members of two networks of organizations focusing on sustainability and corporate social responsibility (Kauri and Positive Entrepreneurs) and as such having a keen interest in non-financial performance. By synthesizing these opinions after each round, the researcher pursues consensus within the panel of experts [60,62,63]. After two rounds, the required consensus was achieved which resulted in the removal of 13 indicators and the selection of 40 indicators. Table 2 gives an overview of the removed and accepted indicators.

Table 2. Overview of the indicators selected through literature review, focus groups and Delphi panel.

Indicators Identified through Literature Review and Focus Groups (53)	Consensus in Delphi Panel (40)
Economic performance	
Literature and approved by focus groups	
Market share in comparison to important competitors	
Growth in market share	
Additional indicators focus groups	
Received subsidies and donations	
Innovativeness	X
Proactiveness	X
Risk Taking	X
Environmental performance	
Literature and approved by focus groups	
Use of renewable energy	X
Transportation of materials and goods	X
Transportation of the members of the organization's workforce	X
Waste reduction	X
Use of sustainable materials	X
Environmental policy	X

Table 2. Cont.

Indicators Identified through Literature Review and Focus Groups (53)	Consensus in Delphi Panel (40)
Community performance	
Literature and approved by focus groups	
Offering job opportunities	
Hiring disadvantaged people	X
Local suppliers	X
Local customers	
Philanthropy	
Partnerships	X
Being responsive to complaints of customers	
Adaptation of products and services to satisfy complaints of customers	
Additional indicators focus groups	
Informing the local community	X
Offering traineeships to students	X
Offering products/services to vulnerable people	X
Addressing unsolved problems in society	X
Human performance	
Literature and approved by focus groups	
Supporting learning initiative	X
Policy on education and training	X
Providing education and training	X
Diversity management	X
Equal opportunities for minorities	X
Involvement of personnel in education and training	X
Age sensitive personnel policy	
Work-life balance	X
Interaction between employees	X
Goal oriented HRM	X
Additional indicators focus groups	
Development/personal growth of personnel	X
Absenteeism through illness	
Support on the work floor	X
Job satisfaction	X
Governance performance	
Literature and approved by focus groups	
Board diversity	
No CEO duality	
Independent board members	X
Adaptation of the composition of the board	X
Clear organizational mission and goals	X
Engagement of board members toward the mission and goals of the organization	X
Involvement of the board in strategic initiatives	X
Clarity of roles (of board members and management team)	X
Participative decision-making	X
Goals meeting the needs of the stakeholders	X
Adaptation to changes in the environment	X
Efficient, well prepared board meetings	
Preparedness to learn from mistakes	X
External communication to stakeholders	X

3.4. Phase 4: Survey Instrument Development and Administration

As explained earlier, next to the selection of relevant performance indicators which can serve the external purpose of reporting to external stakeholders, we aim at developing a measurement instrument that social enterprises can use as an internal, self-assessment tool. Therefore, the 40 selected indicators were concretized in a survey instrument. Questionnaires are the most commonly used method of data collection in field research [33] and, over the past several decades, scales have been developed suitable for assessing input, throughput and output of the performance of organizations.

In the next section (Phase 5), the items and scales used to measure the indicators are discussed for each performance domain separately. To achieve high levels of content validity, most of the constructs and measures used in the instrument were already verified in earlier research. Next, we created measures by adapting existing scales.

Given the purpose of developing a measurement instrument suitable for a broad range of social enterprises, the survey was contributed to different groups of organizations, encompassing the sector of social enterprises in Flanders, Belgium. The work was carried out with the active help of the Flemish government who provided the sample for the study.

The following organizations were selected: (1) sheltered workshops and social workshops: established with the main purpose of reintegrating job seekers who face difficulties to find a job in the regular job market because of physical, social or psychological problems, mainly operating in packaging, assembly, gardening, recycling, and printing [12], (2) local service economy initiatives: social enterprises closely connected to local authorities, offering jobs to long-term unemployed people in combination with offering quality services to the local community and households (e.g. cleaning services, shopping assistance for the elderly) [34], (3) work experience enterprises and work care initiatives, offer a job to long-term unemployed people and are mainly active in health and social care or the cultural sector [34], (4) work integration enterprises receive subsidies in return for employing long-term unemployed jobseekers and integrate them into their regular staff [34] and (5) cooperatives: member-based democratic organizations [64].

The survey was distributed to the top managers of 1018 social enterprises. These top managers have an overview of the overall performance of the organization, including the different performance domains which are part of our measurement instrument. The survey was distributed using a web-based tool (Qualtrics). After a period of intensive follow-up (mail and telephone) of the responses, a total of 244 line managers completed the survey, yielding a response rate of 24%. After removing incomplete surveys, our results are based on the responses from top managers of 241 organizations. The age of the organization in our sample varies between 2 and 93 years old, with an average age of 26. The number of employees ranges from 1 to 2023 with an average of 147 employees. A total of 84% of the organizations are SMEs with less than 250 employees. Table 3 gives an overview of the population of 1018 social enterprises and of the 241 social enterprises that participated.

Table 3. Overview of the population and sample of the survey.

	Population	Sample
Sheltered an Social Workshops	80 (8%)	44 (18%)
Local service economy initiatives	206 (20%)	60 (25%)
Work experience enterprises/Work care initiatives	278 (27%)	81 (34%)
Work integration enterprises	293 (29%)	35 (15%)
Cooperatives	161 (16%)	21 (9%)
Total	1018 (100%)	241 (100%)

3.5. Phase 5: Validation of Relevant Indicators and the Assessment Tool

The aim of this phase in the research process is twofold: (1) reduction of the number of indicators by identifying overarching performance indicators, encompassing different of the retained indicators and (2) validation of the developed survey instrument. Building on scale development and construct validation literature [31,32], we use exploratory and confirmatory factor analysis to reach these goals and we assess the internal consistency of the remaining scales using Cronbach's alpha [65,66].

We will discuss the analyses and results for each performance domain separately (economic, environmental, community, human and governance). First we will give an overview of the items and scales used to measure the selected indicators. When analyzing the data, we first used exploratory factor analysis (EFA). Because some indicators are measured using adapted scales, have not been validated in prior work or are measured using a single item, running a factor analysis for the

items of each indicator separately would be inappropriate. Therefore we conducted, for each performance domain, EFA of all items. If, within a performance domain, items used to measure different performance indicators load on a latent factor, we can reduce the number of indicators by detecting an overarching indicator. Items that load insufficiently onto one factor will be removed if different items are used to measure the indicator [65].

We build on the results of the EFA to specify the factor models used in the confirmatory factor analysis (CFA). We conducted confirmatory factor analysis using the Lavaan package, developed for Structural Equation Modeling (SEM) in the statistical program R [67]. Because we use categorical data (ordinal variables using Likert scales and dichotomous variables), we use the robust weighted least squares (WLSMV) estimator [68]. The Chi Square statistic is commonly reported in CFA research, more specifically, we report the Chi-Square test statistic, divided by the degrees of freedom (c^2/df) [65]. Next to the Chi-Square statistic, it is suggested to take into consideration different other fit indices to evaluate the model fit [66,68]. Brown [68] (p.82) distinguishes three categories of fit indices and advises to report at least one index from each category. Following the advice of Brown (2006), we report the Standardized Root Mean Square Residual (SRMR), the Root Mean Square Error of Approximation (RMSEA), the Comparative Fit Index (CFI) and the Tucker-Lewis Index (TLI).

There is no consensus on the cutoff values that should be used to evaluate model fit. It is even argued that the use of absolute cutoff values is inadvisable because fit indices are influenced by different aspects of the research setting, e.g. sample size and type of data [66,68]. However, there are some guidelines for the fit indices we use in our study. For the c^2/df ratio, Janssens, Wijnen, De Pelsmacker and Van Kenhove [65] mention as criterion < 2 , while Hair, Black, Babin and Anderson [66] mention < 3 . For the SRMR, Hu and Bentler [69] use a cutoff value of 0.08, while Hair, Black, Babin and Anderson [66] mention that an SRMR over 0.1 suggests a problem with fit. Concerning the RMSEA, the cutoff value of 0.06 proposed by Hu and Bentler [69] is often referred to. However, Brown [68] mentions that RMSEAs in the range of 0.8–0.1 suggest a mediocre fit and that models with RMSEA over 0.1 should be rejected. For CFI and TLI, Hu and Bentler [69] suggest values ≥ 0.95 , but different authors indicate that values in the range 0.9–0.95 indicate acceptable fit [65,66,68].

Finally, we assess the internal consistency or reliability of the scales used for measuring the different indicators, by reporting the Cronbach's alpha [65]. Based on Hair, Black, Babin and Anderson [66], a value above 0.7 is considered to indicate a strong reliability, while a value above 0.6 indicating a satisfactory reliability, allowing the use of summated scales.

3.5.1. Economic Performance

Economic performance is related to conditions supporting the financial sustainability of organizations. As mentioned earlier, the focus is not on traditional financial indicators such as profit, cash flow, Return on Assets (ROA) and Return on Investment (ROI) [45], but on indicators positively influencing these indicators. Based on the focus group sessions and the Delphi panel, three indicators are selected to assess the economic performance of social enterprises. Table 4 gives an overview of the indicators, items and scales used to evaluate economic performance.

These indicators (innovation, proactiveness and risk taking) are related to the entrepreneurial orientation of organizations. Therefore, we used the measure introduced by Helm and Andersson [70], specifically developed to evaluate the entrepreneurial orientation of social enterprises and comprising three subscales to measure innovation, proactiveness and risk taking. The scale measures along a continuum: two opposite statements are formulated and respondents are asked to indicate on an 8-point Likert scale which statement best characterizes their organization.

Table 4. Economic performance: overview of indicators, items and scales.

Innovation	
<i>Presently and during the last five years my organization has:</i>	
ECON1*	Placed a strong emphasis on the maintenance of tried-and-true products or services– Placed a strong emphasis on the development of new products or services
ECON2	Placed a strong emphasis on the maintenance of established organizational processes– Placed a strong emphasis on the development of new organizational processes
ECON3	Introduced no new processes, policies, products or services– Introduced many new processes, policies, products and services
ECON4	Made only minor changes in processes, policies, products or services– Made major changes in processes, policies, products or services
Proactiveness	
<i>Presently and during the last five years my organization:</i>	
ECON5	Is very seldom the first organization to introduce new products/services, administrative techniques, operating technologies, etc.– Is very often the first organization to introduce new products/services, administrative techniques, operating technologies, etc.
ECON6	Been reticent to exploit changes in the field—Exploited changes in the field
ECON7	Followed the lead of similar service providers—Provided the lead for similar service providers
Risk Taking	
<i>Presently and during the last five years my organization:</i>	
ECON8	Conducted itself consistently with the behavioral norms of the operating environment, industry or sector– Conducted itself in conflict with the behavioral norms of the operating environment, industry or sector
ECON9	Selected projects that support the organization's public image– Selected projects that may alter the organization's public image
ECON10	Made decisions that maintain staff stability– Made decisions that created changes in staff stability

Measured on an 8-point Likert scale. Based on Helm and Andersson (2010) [70]; * Item removed after EFA.

We conducted a principal component exploratory factor analysis using varimax rotation of the 10 items of the scale. The results are reported in Table 5. The results show three factors with eigenvalues greater than one, explaining 72% of the variance, corresponding to the subscales identified by Helm and Andersson (2010). Item “ECON1” has a high factor loading (>0.5) on “Innovation” as well as on “Proactiveness”. For that reason we decided to exclude “ECON1” in the confirmatory factor analysis. The other items loaded sufficiently onto one single (expected) factor.

Table 5. Economic Performance: items and item loadings EFA.

	ECO1 Innovation	ECO2 Proactiveness	ECO3 Risk Taking
ECON1	0.561	0.505	0.304
ECON2	0.812	0.122	0.265
ECON3	0.803	0.249	0.045
ECON4	0.806	0.350	0.094
ECON5	0.226	0.832	0.195
ECON6	0.333	0.805	0.044
ECON7	0.179	0.822	0.225
ECON8	0.092	0.320	0.745
ECON9	0.036	0.129	0.824
ECON10	0.327	0.029	0.757

Principal component factor analysis, varimax rotation.

In a next step, we used CFA to test a second order model. Specifically, we checked whether the results of EFA are confirmed by CFA and whether the three detected factors (innovation, proactiveness and risk taking) load onto the second order factor “Economic Performance”. The results are reported in Figure 2.

The results show acceptable fit indices, with $c^2/df = 1.99$; CFI = 0.936; TLI = 0.905; RMSEA = 0.064 and SRMR = 0.043. All factor loadings are significant ($p < 0.001$) and indicate strong factor loadings. The results indicate that the indicators (1) innovation, (2) proactiveness and (3) risk taking are relevant to measure the economic performance of social enterprises. Finally, we checked the internal consistency of the scales used to measure these indicators by calculating Cronbach’s alpha. The results reported in Figure 2, indicate strong scale reliability (>0.70): innovation ($\alpha = 0.829$), proactiveness ($\alpha = 0.848$), risk taking ($\alpha = 0.739$).

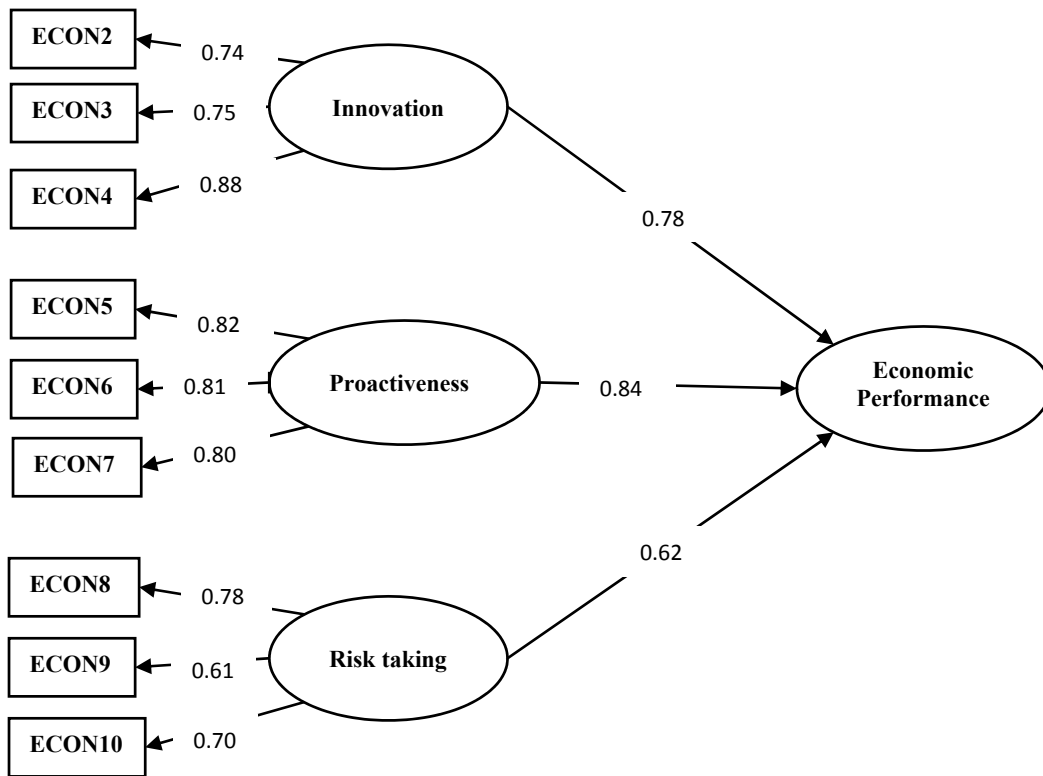


Figure 2. Economic Performance: items and item loadings CFA. Standardized item loadings using the WLSMV estimator: $P < 0.001$ for all loadings; $c^2/df = 1.99$; CFI = 0.936; TLI = 0.905; RMSEA = 0.064; SRMR = 0.043.

3.5.2. Environmental Performance

Environmental performance refers to the efforts organizations make to protect nature [47]. Based on the focus group sessions and the Delphi panel, seven indicators are selected to assess the environmental performance of social enterprises. Table 6 gives an overview of the indicators, items and scales used to evaluate the environmental performance.

Table 6. Environmental performance: overview of indicators, items and scales.

Transportation of Materials and Goods	
ENV1	Our organization deliberately selects cleaner transportation methods for materials and goods Mishra and Suar (2010) [71]
Transportation of the members of the organization's workforce	
ENV2	Our organization encourages employees to use ecological transportation modes Adaptation based on GRI [27]
Use of sustainable materials	
ENV3	Our organization uses recycled input materials Adaptation based on GRI [27]
ENV4	Our organization takes the initiative to use environmental-friendly natural resources Chen <i>et al.</i> (2008) [72]
ENV5	Our organization has a preference for green products in purchasing Mishra and Suar (2010) [71]
ENV6*	Our organization has implemented sustainability criteria for the procurement of goods and services Adaptation based on GRI [27]
Environmental policy	
ENV7	Our organization has incorporated environmental performance objectives in organizational plans Rettab <i>et al.</i> (2009) [73]
ENV8*	Our organization is concerned about the protection of the natural environment Adaptation based on GRI [27]
ENV9	Our organization has a clear environmental policy Mishra and Suar (2010) [71]
Waste reduction	
ENV10	Our organization has reduced the amount of waste in recent years Adaptation based on GRI [27]
Environmental performance measurement	
ENV11	Does your organization measure the organization's environmental performance? Rettab <i>et al.</i> (2009) [73]
Use of renewable energy	
ENV12	Does your organization use energy produced from renewable sources? O'Connor & Spangenberg (2008) [74]

* Item removed after EFA; All items measured on a 7-point Likert scale except ENV11 (yes/no) and ENV12 (yes/no).

We conducted a principal component exploratory factor analysis using varimax rotation of the 10 items measured using a 7-point Likert scale. The results are reported in Table 7. Based on the results of the EFA, we can distinguish three factors, explaining 74% of the variance. These three factors are related to (1) Transportation, (2) Use of ecological materials and (3) Environmental performance management. Item "ENV6" has a high factor loading (>0.5) on "Ecological Materials" as well as on "Environmental performance management". For that reason, we decided to exclude "ENV6" in the confirmatory factor analysis. Item "ENV8" seems to be scattered across the three factors and does not load sufficiently onto one single factor. For that reason also "ENV8" is eliminated in the confirmatory factor analysis. The other items load sufficiently onto one single factor.

Table 7. Environmental Performance: items and item loadings EFA.

	EN1 Transportation	EN2 Ecological Materials	EN3 Environmental Performance Management
ENV1	0.757	0.186	0.344
ENV2	0.832	0.286	0.120
ENV3	0.264	0.747	0.301
ENV4	0.215	0.843	0.292
ENV5	0.278	0.790	0.257
ENV6	0.004	0.568	0.626
ENV7	0.274	0.400	0.717
ENV8	0.408	0.548	0.347
ENV9	0.231	0.366	0.812
ENV10	0.283	0.168	0.766

Principal component factor analysis, varimax rotation.

We checked whether the results of EFA are confirmed by CFA and whether the three detected factors (transportation, ecological materials and environmental performance management) load onto the second order factor “Environmental Performance”. We also added the two items measured as dummy variables. Based on content, we added ENV11 to the factor “environmental performance management” and ENV12 to the factor “ecological materials”.

The fit indices reveal an acceptable fit, except for TLI: $c^2/df = 2.79$; CFI = 0.922; TLI = 0.890; RMSEA = 0.086 and SRMR = 0.047. All factor loadings are significant ($p < 0.001$), but the factor loading of ENV12 is low (0.283). Further investigation of the model in Lavaan shows that the fit of the model will be better if ENV12 is moved to the factor “environmental performance management”. This is acceptable as the use of renewable energy can be considered as an environmental result. This is comparable to the fact that ENV10 (waste reduction), based on the results of EFA, also loads on this factor.

The fit of this model is better. Figure 3 gives an overview of the results. All fit indices show an acceptable fit: $c^2/df = 1.73$; CFI = 0.968; TLI = 0.955; RMSEA = 0.055 and SRMR = 0.036. All factor loadings are significant, however the factor loading of ENV12 is still low (0.304). Finally, we checked the reliability of the scales used to measure the three remaining indicators by calculating Cronbach’s alpha. The results indicate a satisfactory scale reliability (>0.60): Transportation ($\alpha = 0.696$), Ecological materials ($\alpha = 0.877$), Environmental Performance Management ($\alpha = 0.829$). The results indicate that the originally selected seven indicators can be reduced to three indicators, relevant for measuring the environmental performance of social enterprises.

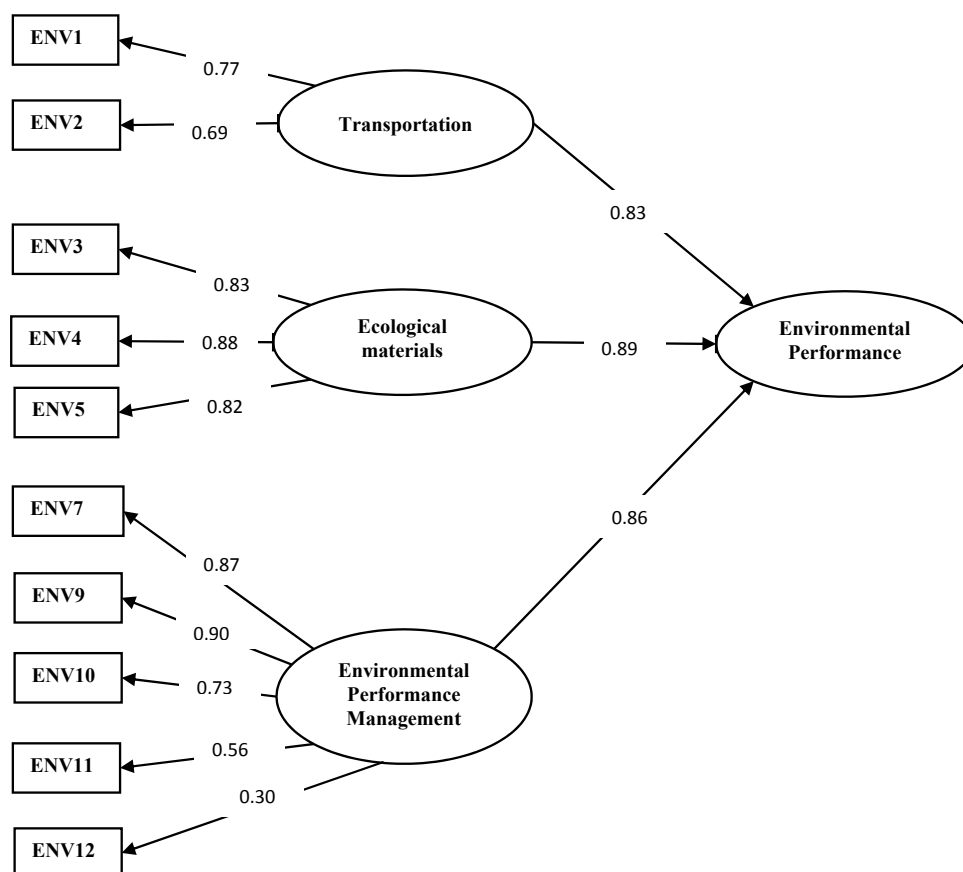


Figure 3. Environmental Performance: items and item loadings CFA. Standardized item loadings using the WLSMV estimator: $p < 0.001$ for all loadings; $c2/df = 1.73$; CFI = 0.968; TLI = 0.955; RMSEA = 0.055; SRMR = 0.036.

3.5.3. Community Performance

Community performance refers to how organizations deal with their responsibilities in society [48]. Following the results of the focus groups and the Delphi panel, seven indicators are selected. Table 8 gives an overview of the indicators, items and scales used to evaluate the community performance.

Table 8. Community performance: overview of indicators, items and scales.

Hiring disadvantaged people	
COM1	Our organization actively hires immigrants Adaptation based on Graafland <i>et al.</i> (2004) [75]
COM2	Our organization actively hires low skilled people Adaptation based on Graafland <i>et al.</i> (2004) [75]
COM3	Our organization actively hires elderly people Adaptation based on Graafland <i>et al.</i> (2004) [75]
Informing the local community	
COM4	Our organization informs the local community by organizing presentations, company visits Adaptation based on CAF [76]
Offering traineeships to students	
COM5	Our organization offers traineeships to students Adaptation based on CAF [76]

Table 8. Cont.

Offering products/services to vulnerable people	
COM6	Our organization offers products and/or services to vulnerable people Adaptation based on CAF [76]
Addressing unsolved problems in society	
COM7	Our organization addresses unsolved societal problems Adaptation based on CAF [76]
Partnerships	
COM8	Our organization pursues partnerships with: <ul style="list-style-type: none"> • Governments • For profit organizations • Social economy organizations • Labor agencies • Other community organizations Adaptation based on Mishra and Suar (2010) [71]
Local suppliers	
COM9 *	Our organization mainly has local (Flemish) or regional (Belgian) suppliers Adaptation based on GRI [27]

* Item removed after CFA; All items measured on a 7-point Likert scale except COM8 (sum of different kind of partnerships) and COM9 (yes/no).

We conducted a principal component exploratory factor analysis using varimax rotation of eight items. As we did in the analyses of environmental performance, we did not take in consideration COM9 because it is a dummy variable. The results are reported in Table 9. Based on the results of the EFA, we can distinguish two factors, explaining 56% of the variance. The first factor is related to the indicator “Hiring disadvantaged people”. All the other indicators load on a second factor, that we call “Community responsibilities”.

The results of the EFA reveal that all indicators of community performance load onto one factor, except for the items related to the indicator “hiring disadvantaged people”. This indicates that hiring disadvantaged people is considered as a distinctive performance indicator of community performance in relationship to the other indicators. We use CFA to check if this distinction is confirmed. Furthermore, using CFA, we want to investigate whether the two detected factors load on the second order construct “community performance”. In this stage we also add COM9, measured as a dummy variable. In our model we add COM9 to the factor “Community responsibilities” as choosing local suppliers is not related to the hiring of disadvantaged people.

Table 9. Community Performance: item and item loadings EFA.

	C1	C2
	Hiring Disadvantaged People	Community Responsibilities
COM1	0.887	0.059
COM2	0.727	0.092
COM3	0.845	0.041
COM4	−0.040	0.695
COM5	0.200	0.459
COM6	0.126	0.766
COM7	0.057	0.848
COM8	−0.013	0.639

Principal component factor analysis, varimax rotation.

Testing this second order model in Lavaan, the fit indices show an acceptable fit ($c^2/df = 1.69$; CFI = 0.932; TLI = 0.905; RMSEA = 0.054 and SRMR = 0.055), however the factor loading of COM9 (local suppliers) is low (0.173) and not significant ($p < 0.1$). A possible explanation is that the organizations in our sample are mainly small, locally embedded organizations. A total of 89% of the social enterprises in our sample mainly have local or regional suppliers. As such, having local suppliers is not a distinguishing factor in our sample. Because of the low, insignificant factor loading, we decided to exclude COM9.

Figure 4 gives an overview of the CFA results of the adapted model. The fit indices show a good fit: $c^2/df = 1.78$; CFI = 0.943; TLI = 0.916; RMSEA = 0.057 and SRMR = 0.053. All factor loadings are significant ($p < 0.001$). The Cronbach’s alpha of the scales used to measure “hiring disadvantaged people” ($\alpha = 0.767$) and “community responsibilities” ($\alpha = 0.718$) reveal a strong reliability. The results indicate that the seven indicators selected for assessing the community performance of social enterprises can be reduced to two indicators.

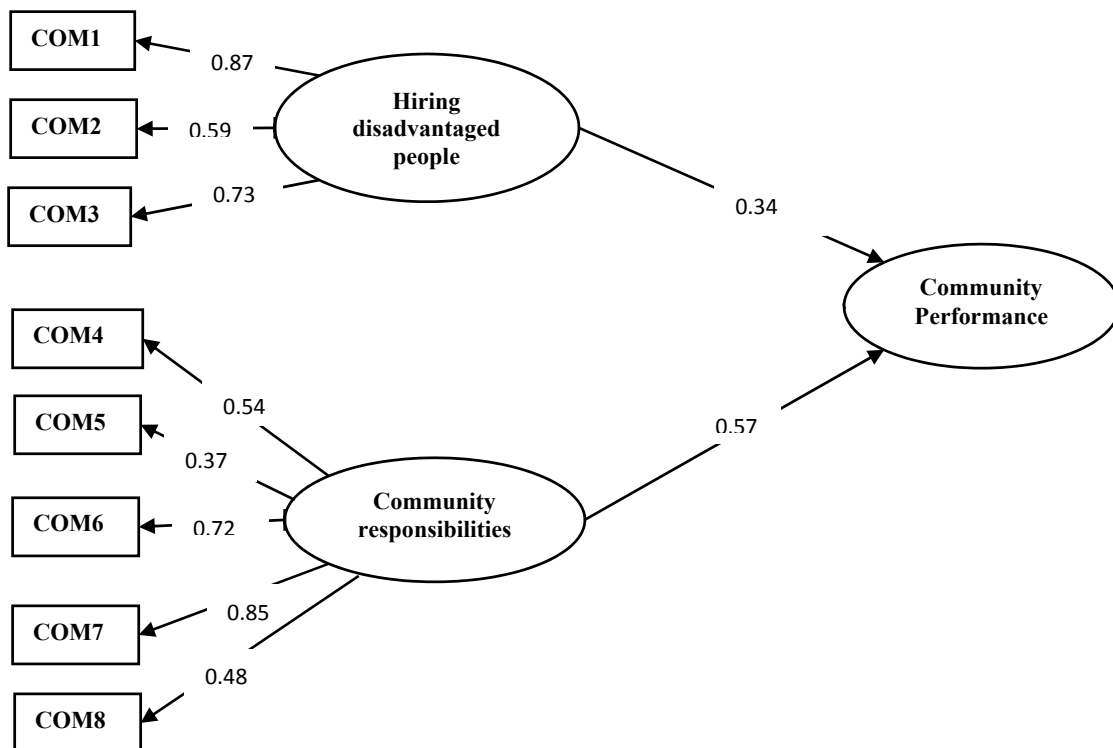


Figure 4. Community Performance: item and item loadings CFA. Standardized item loadings using the WLSMV estimator; $p < 0.001$ for all loadings; $c^2/df = 1.78$; CFI = 0.943; TLI = 0.916; RMSEA = 0.057; SRMR = 0.053.

3.5.4. Human Performance

Human performance refers to the relationship of the organization with its workforce [46]. Based on the results of the focus groups and the Delphi panel, 12 indicators are selected. Table 10 gives an overview of the indicators, items and scales used to evaluate the human performance.

Table 10. Human performance: overview of indicators, items and scales.

Providing education and training Calantone <i>et al.</i> (2002) [77]	
HUM1	Our organization has a strong ability to learn and this offers us a competitive advantage
HUM2	The basic values of this organization include learning as key to improvement
HUM3	The sense around here is that employee learning is an investment, not an expense
HUM4	Learning in my organization is seen as a key commodity necessary to guarantee organizational survival
Development/ personal growth of employees	
HUM5	We develop our employees aiming at job rotation within our organization Adaptation based on GRI [27]
Supporting learning initiative	
HUM6	Our organization supports all employees who want to pursue further education Rettab <i>et al.</i> (2009) [73]
Equal opportunities for minorities	
HUM7	Our organizations has a policy concerning equal rights and non-discrimination O'Connor & Spangenberg (2008) [74]
Involvement of personnel in education and training	
HUM8	Our organization involves the employees in the planning of education and training Adaptation based on CAF [76]
Interaction between employees	
HUM9	We pay attention to good relationships between our employees Adaptation based on ISO26000 [78]
Goal oriented HRM	
HUM10	Our HR-policy is carefully planned Adaptation based on GRI [27]
HUM11	Our HR-policy is carefully evaluated Adaptation based on GRI [27]
Job satisfaction	
HUM12	Our organization pays attention to individual job satisfaction Adaptation based on GRI [27]
Diversity management	
HUM13	Our organization has a policy on diversity management Cuesta Gonzalez <i>et al.</i> (2006) [79]
Policy on education and training	
HUM14	Our organization has a policy for the training and development of employees Mishra and Suar (2010) [71]
Support on the work floor Adaptation based on Heslin <i>et al.</i> (2006) [80]	
HUM15*	We support our employees in taking on new challenges
HUM16	We offer useful suggestions regarding how employees can improve their performance
HUM17	We provide constructive feedback to employees regarding areas for improvement
HUM18	We help employees to analyze their performance
HUM19	We provide guidance regarding performance expectations
Work-life balance	
HUM20	Our organization is successful in balancing paid work and family life Adaptation based on Milkie & Peltola (1999) [81]

* Item removed after EFA; All items measured on a 7-point Likert scale.

The results of the EFA are shown in Table 11. Based on the results of EFA, we can distinguish four factors, which we identify as (1) performance support, (2) training & development, (3) HR-policy and (4) diversity management. The results are not straightforward: different items have high factor loadings on different factors. A second problem is that in some cases the loading on a specific factor cannot be explained based on the content of the item. This is not exceptional in EFA. Janssens, Wijnen, De Pelsmacker and Van Kenhove [65] suggest that content should be taken into consideration and that content takes precedence over factor loadings.

Table 11. Human Performance: items and item loadings EFA.

	H1 Performance Support	H2 Training & Development	H3 HR-Policy	H4 Diversity Management
HUM1	0.263	0.650	0.301	0.142
HUM2	0.318	0.835	0.207	0.111
HUM3	0.287	0.718	0.151	0.352
HUM4	0.235	0.770	0.303	0.116
HUM8	0.318	0.729	0.204	0.122
HUM16	0.805	0.362	0.199	0.096
HUM17	0.805	0.338	0.191	0.157
HUM18	0.786	0.278	0.247	0.140
HUM19	0.814	0.244	0.186	0.137
HUM10	0.239	0.319	0.758	0.256
HUM11	0.274	0.340	0.746	0.239
HUM7	0.144	0.317	0.309	0.772
HUM13	0.160	0.266	0.367	0.743
HUM15	<u>0.686</u>	0.299	<u>0.400</u>	0.036
HUM5	0.255	<u>0.413</u>	<u>0.418</u>	0.226
HUM6	<u>0.604</u>	<u>0.377</u>	0.112	0.324
HUM14	<u>0.582</u>	0.243	<u>0.472</u>	0.128
HUM9	0.291	<u>0.663</u>	0.092	<u>0.390</u>
HUM20	<u>0.593</u>	0.093	−0.109	<u>0.514</u>
HUM12	0.261	<u>0.545</u>	0.312	0.243

Principal component factor analysis, varimax rotation.

In analyzing and evaluating the factor loadings we took different steps. In a first step, we look for items with significant factor loadings on different factors. Based on our sample size, we consider values above 0.375 as significant factor loadings [65]. If there are different items used to measure the indicator, we remove the item. This is the case for HUM15. If the indicator is measured using a single item, we decide not to remove the item because this would imply removing the indicator. Instead, we evaluate based on content to which factor we will add the item in the CFA-model. This is the case for HUM5, HUM6, and HUM14. For Item “HUM5” (significant loading on H2 and H3), we decide to assign it to H2 as HUM5 is more related to education and training. Item “HUM6” has a significant factor loading on H1 and H2. Based on the content, we decide to add it to H2. Item “HUM14” has a significant loading on H1 and H3. As it is more related to HR-policy (H3) than to Performance Support (H1), we assign it to H3. For some items with significant factor loadings on different factors, we notice that it is difficult to assign them to one of the factors. This is the case for HUM9 and HUM20. Therefore, we decide to assess these items separately and not assigning them to a first order factor. Instead, we will add these items to the model, directly loading on the second order factor “Human Performance”.

In a second step, we check if the items with only one significant factor loading can be assigned to that factor based on content. HUM12 has a significant factor loading on H3 but is not related to training and development. Neither is it related to one of the other factors. Therefore, we decide, similarly to HUM9 and HUM20, to load it directly onto the second order factor Human Performance.

We use CFA to test this second order model. The results of the second order CFA model show acceptable fit indices: $c^2/df = 2.94$; CFI = 0.916; TLI = 0.903; RMSEA = 0.09 and SRMR = 0.066. All factor loadings are significant. The fit of the model will however be better when HUM14 is removed to H2. Based on content, HUM14 is as well related to H2 (Training and Development) as to H3 (HR-policy). However because the overall fit is better when HUM14 is assigned to H2, we decide to assign HUM14 to H2. Figure 5 gives an overview of the results of the adapted model. The fit indices show a good fit: $c^2/df = 2.9$; CFI = 0.918; TLI = 0.905; RMSEA = 0.089 and SRMR = 0.058. All factor loadings are significant ($p < 0.001$). Finally, we checked the reliability of the scales used to measure the four remaining indicators by calculating Cronbach’s alpha. The results indicate a strong scale reliability (>0.70): Performance support ($\alpha = 0.938$), Training & Development ($\alpha = 0.907$), HR-policy ($\alpha = 0.883$) and Diversity Management ($\alpha = 0.843$). The results indicate that the originally selected 12 indicators can be reduced to seven indicators (the four indicators H1, H2, H3 and H4 and the single-item indicators HUM9, HUM12 and HUM20), relevant for measuring the human performance of social enterprises.

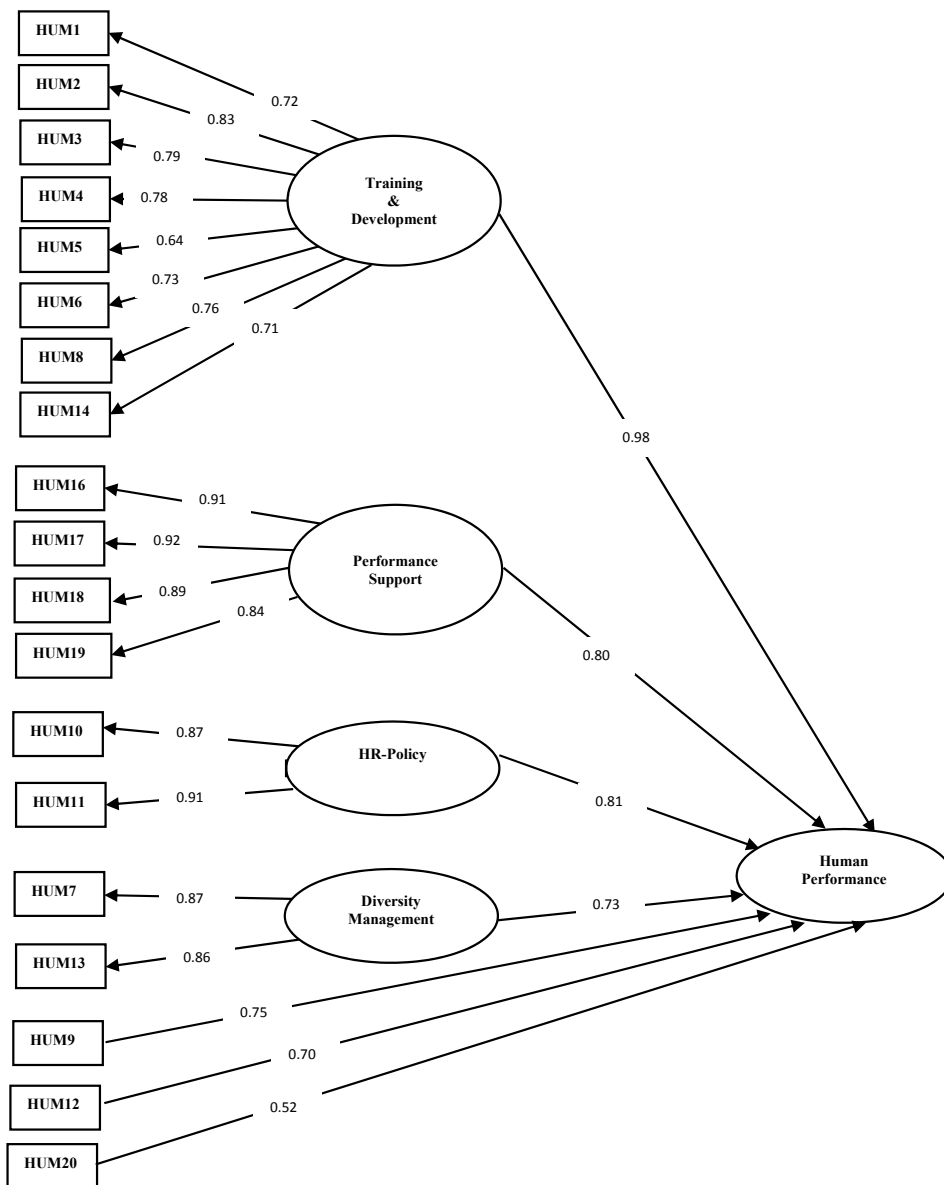


Figure 5. Human Performance: items and item loadings CFA. Standardized item loadings using the WLSMV estimator: $P < 0.001$ for all loadings; $c^2/df = 2.9$; CFI = 0.918; TLI = 0.905; RMSEA = 0.089; SRMR = 0.058.

3.5.5. Governance Performance

Governance performance focuses on good governance practices. On the one hand, it is related to best practices regarding board composition and board practices. On the other hand, it refers to having clear organizational goals taking into consideration stakeholder expectations [49,50].

Based on the results of the focus groups and the Delphi panel, 11 indicators are selected. Table 12 gives an overview of the indicators, items and scales used to assess governance performance.

Table 12. Governance performance: overview of indicators, items and scales.

Adaptation of the composition of the board	
GOV1	New board member are selected to meet the organization's changing needs Adaptation based on Herman and Renz (2004) [82]
Adaptation to changes in the environment Adaptation based on Jackson and Holland (1998) [83]	
GOV2	The board of directors is able to cope with changes in the legal environment.
GOV3	The board of directors is able to cope with changes in the economic environment.
GOV4	The board of directors is able to cope with changes in the political environment.
GOV5	The board of directors is able to cope with changes in the needs of stakeholders.
Engagement of board members toward the mission and vision of the organization Fredette and Bradshaw (2012) [84]	
GOV6	Board members share the same ambitions and vision for the organization.
GOV7	Board members enthusiastically pursue collective goals and mission.
GOV8	Board members are committed to the goals of the organization.
GOV9	Board members view themselves as partners in charting the organization direction.
GOV10	There is a commonality of purpose among board members of this organization.
GOV11	Everyone in the board of directors is in total agreement on our organization's vision.
Participative decision-making Li and Hambrick (2005) [85]	
GOV12	All the board members have a voice in major decisions.
GOV13	Communications among board members can best be described as open and fluid.
GOV14	When major decisions are made, board members collectively exchange their points of view.
GOV15	Board members frequently share their experience and expertise.
Clarity of roles Gill <i>et al.</i> (2005) [86]	
GOV16	Board members demonstrate clear understanding of the respective roles of the board and CEO
Preparedness to learn from mistakes Jackson and Holland (1998) [83]	
GOV17	In the board of directors we discuss about what we can learn from a mistake we have made
External communication to stakeholders Jackson and Holland (1998) [83]	
GOV18	This board communicates its decisions to everyone who is affected by them
GOV19	The board is actively involved in long-term strategic decision-making
GOV20	The board is actively involved in implementing long-term strategic decision-making
GOV21	The board is actively involved in promoting strategic initiatives
Goals meeting the needs of the stakeholders Rettab <i>et al.</i> (2009) [73]	
GOV22	The goals of our organization meet the needs and requests of all our stakeholders
Clear organizational mission and goals Wright (2007) [88]	
GOV23	It is easy to explain the goals of this organization to outsiders
GOV24	This organization's mission is clear to everyone who works here
GOV25	This organization has clearly defined goals
Independent board members Hillman <i>et al.</i> (2000) [89], Haynes and Hillman (2010) [90]	
GOV26 *	Does the organization has outside, independent directors?

* Item removed after CFA; All items measured on a 7-point Likert scale, except for GOV26 (Yes/No).

We conducted a principal component exploratory factor analysis using varimax rotation of the 25 items measured on a 7-point Likert scale. The results are reported in Table 13 and reveal five factors, explaining 70% of the variance. However, the results are not straightforward: several items have significant factor loadings on different items and sometimes the loading on a factor is not in line with the content of the item, making a thorough evaluation necessary.

Table 13. Governance Performance: items and item loadings EFA.

	G1 Shared Vision	G2 Adaptability	G3 Strategic Board Role	G4 Participative Decision-Making	G5 Clear Organizational Goals
GOV1	0.127	0.553	0.140	0.071	0.003
GOV2	0.184	0.834	0.210	0.184	0.077
GOV3	0.210	0.764	0.315	0.082	0.158
GOV4	0.174	0.789	0.044	0.099	0.156
GOV5	0.237	0.771	0.236	0.161	0.166
GOV6	0.781	0.259	0.171	0.191	0.095
GOV7	0.799	0.232	0.108	0.121	0.162
GOV8	0.761	0.154	0.176	0.293	0.228
GOV9	0.784	0.309	0.200	0.231	0.045
GOV10	0.876	0.167	0.060	0.105	0.123
GOV11	0.874	0.215	0.083	0.177	0.111
GOV12	0.402	0.143	0.154	0.567	0.076
GOV13	0.656	0.320	0.072	0.454	0.018
GOV14	0.369	0.265	0.070	0.658	0.042
GOV15	0.421	0.279	0.210	0.569	−0.002
GOV16	0.280	0.722	0.103	0.130	0.098
GOV17	0.286	0.554	0.212	0.454	0.042
GOV18	0.247	0.501	0.007	0.482	0.099
GOV19	0.277	0.302	0.693	0.328	0.141
GOV20	0.150	0.264	0.866	0.064	0.078
GOV21	0.140	0.253	0.889	0.093	0.103
GOV22	0.005	−0.041	0.134	0.508	0.540
GOV23	0.014	0.083	0.066	0.036	0.690
GOV24	0.288	0.145	0.048	−0.025	0.810
GOV25	0.169	0.188	0.065	0.066	0.769

Principal component factor analysis, varimax rotation.

First, we look for items with significant factor loadings on different factors. If different items are used to measure the indicator, we consider excluding the item. This is the case for GOV12, GOV13 and GOV15, items used to measure the indicator “Participative decision-making”, but which is apparently closely related to the indicator “Shared vision”. Eliminating these three items implies that “Participative decision-making” would be measured using only one item. We do not consider this as a suitable solution. An alternative solution is to combine “Shared vision” and “Participative decision-making” and exclude GOV14. However, we do not consider this a good solution either as having a shared vision on the mission and goals of the organization is clearly distinct from how issues are discussed within the board and from how decisions are made by the board. Therefore, we decide to keep both the factors G1 and G4 and to keep the items GOV12, GOV13 and GOV15.

Also, GOV17, GOV18 and GOV22 have high factor loadings on different factors. However, because these items are used to measure indicators with a single item, we decide not to remove the items. Instead, we determine to which factor the items can be added based on content. GOV17 has a high factor loading on G2 and G4. Examining content we decide to add GOV17 to G2 because this factor is related to the

extent that boards are able to adapt to changes in the external environment. We consider “preparedness to learn from mistakes” (GOV17) as related to the adaptability of the board of the organization.

GOV18 has a high factor loading on G2 and G4. The indicator is related to communication to the stakeholders, which is not related to having a shared vision within the board nor to participative decision-making in the board. Therefore, we keep it as a separate indicator in the CFA model directly loading on the second order construct “Governance performance”.

GOV22 has a high factor loading on G4 and G5. As the indicator “Goals meeting the needs of the stakeholders” is measured using this single item, we decide not to exclude GOV22. Based on content, we decide to add it to the factor G5 “Clear organizational goals” in the CFA model.

In a next step, we use CFA to test a second order model. The results are reported in Figure 6.

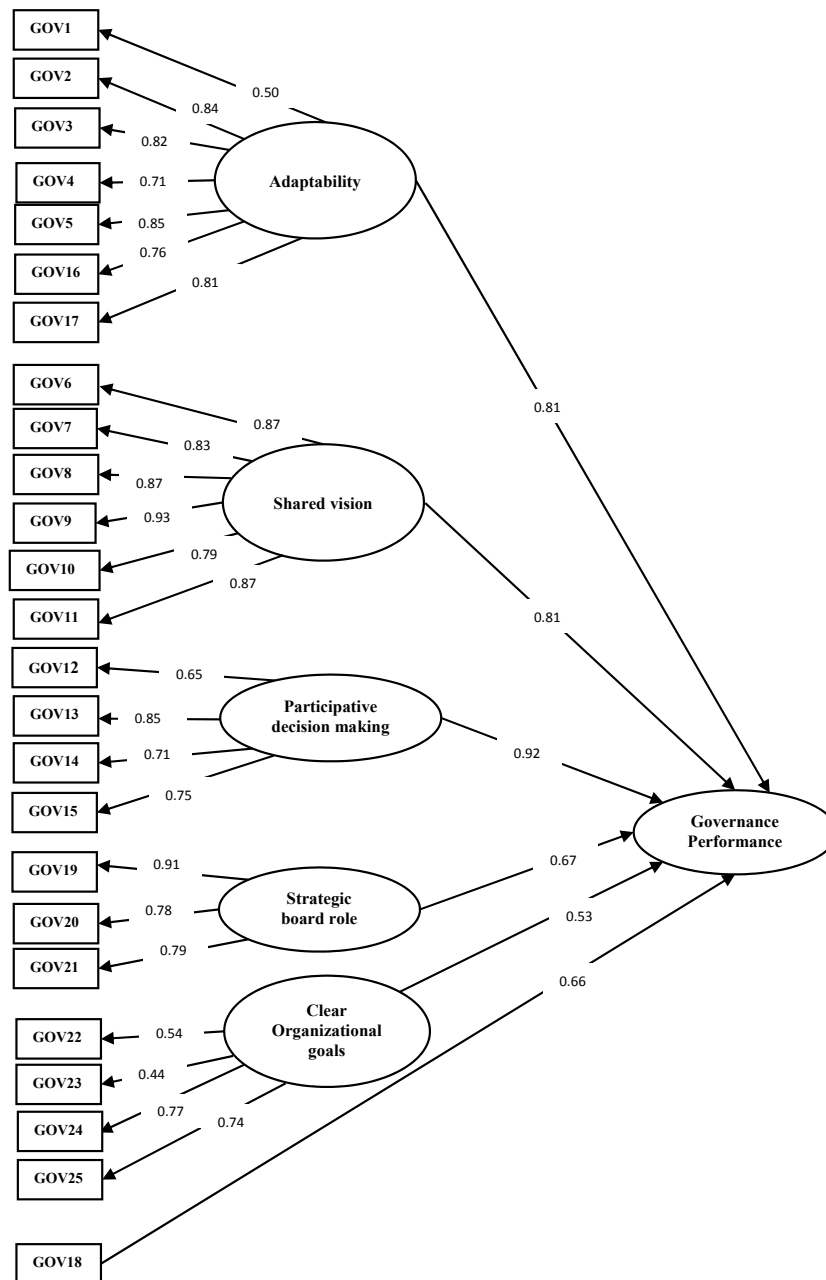


Figure 6. Governance Performance: item and item loadings CFA. Standardized item loadings using the WLSMV estimator: $P < 0.001$ for all loadings; $c^2/df = 1.24$; CFI = 0.942; TLI = 0.935; RMSEA = 0.036; SRMR = 0.057.

Specifically, we test whether the model reflecting the results of EFA and reflecting our interpretation based on content are confirmed by CFA. We also added GOV26, measuring the independent directors as a dummy variable. The results reveal an acceptable fit, but GOV26 has a very low factor loading -0.190 ($p < 0.01$). A possible explanation is that 91% of the organizations in our sample has external directors and as such this indicator is not a distinguishing factor in our sample. Therefore, we decided to remove GOV26 in our analyses.

The results of the adapted model (Figure 6) show acceptable fit, with $c^2/df = 1.24$; CFI = 0.942; TLI = 0.935; RMSEA = 0.036; SRMR = 0.057. All factor loadings are significant ($p < 0.001$) and indicate strong factor loadings. Finally, we checked the reliability of the scales by calculating Cronbach's alpha. The results reported in Figure 6, indicate a strong scale reliability (>0.70): Adaptability ($\alpha = 0.897$), Shared vision ($\alpha = 0.944$), Strategic board role ($\alpha = 0.886$), Participative decision-making ($\alpha = 0.828$) and Clear organizational goals ($\alpha = 0.748$). The results indicate that the 11 originally selected indicators can be reduced to six indicators (G1, G2, G3, G4, G5 and GOV18), relevant for measuring the governance performance of social enterprises.

3.5.6. Overview of Selected Indicators Based on the Results of the Exploratory and Confirmatory Factor Analyses

As result of the exploratory and confirmatory factor analyses, we retain 21 indicators. Table 14 gives an overview of these indicators, for each performance domain.

Table 14. Retained indicators after EFA and CFA.

Economic performance	
Innovation	Proactiveness
Risk taking	
Environmental performance	
Transportation	Environmental performance management
Ecological materials	
Community performance	
Hiring disadvantaged people	Community responsibilities
Human performance	
Performance support	Interaction between employees
Training & Development	Job satisfaction
HR-policy	Work-life balance
Diversity management	
Governance performance	
Shared vision	Participative decision-making
Adaptability of the board	Clear organizational goals
Strategic board role	External communication to stakeholders

4. Discussion and Conclusion

This paper aimed at describing an assessment tool for the organizational performance of social enterprises, as well as reporting on its development and reliability. By developing a set of relevant indicators suitable for external reporting and an internal assessment tool, we answer the calls for a scale development of performance measurement tools that can be implemented in social enterprises. The tool emphasizes the importance of assessment and the use of a valid tool to determine organizational performance in social enterprises.

The results of our study will be discussed in relation to four key notions—robustness, utility, understanding, and relevance—with regard to scale development [91]. Robustness refers to the psychometric qualities of the instrument. With regard to this quality aspect, we conducted exploratory

and confirmatory analysis to identify relevant indicators and we controlled for the internal consistency (or reliability) of the scales used to measure the indicators. Utility relates to the application of an instrument and the implications of the results: the evaluation tool can be regarded as an adequate tool to integrate multiple values of social enterprises. Thirdly, understanding refers to how we should correctly assess and interpret the construct. Since the evaluation tool is carried out in close cooperation and dialogue with stakeholders (managers of social enterprises and experts in the field of social entrepreneurship), we integrated best practices and knowledge in the field based on different perspectives. With regard to relevance, the application of the assessment tool for measuring social performance broadens our view and stimulates us to think beyond financial performance in social enterprises by placing emphasis on different performance dimensions. We emphasize the rigorous process which we have undertaken in the development of this assessment tool. A profound literature study guarantees content validity of the items. Recommendations of experts and stakeholders—in order to adapt the measures to the context of social enterprises—were taken into account during the development of the assessment tool. A major contribution is the sample size and representativeness of the respondents of the scale used in the final validation of the research. We can also stress the innovation in the methodology in this specific context.

We contribute to the literature surrounding performance measurement in social enterprises as it adds to our understanding of the use of performance measurement systems in social enterprises. Existing literature has shown that no large-scale empirical research has been systematically conducted dealing with performance measurement in social enterprises and that the developed tools are either too general or too specific in their design, having an impact on the practical usefulness of these tools. In order to bridge this gap, we used qualitative and quantitative research methods, incorporating the expertise and knowledge of multiple actors in the field of social entrepreneurship, to develop a performance measurement tool, suitable for a wide range of social enterprises.

Moreover, we add to the literature surrounding performance measurement in social enterprises by contributing to the use of performance indicators and performance measurement in the particular context of social enterprises. We acknowledge the role of diverse performance dimensions in this instrument and we take a more comprehensive view of performance than mere financial organizational performance, proposing a much broader concept that encompasses a variety of performance indicators. Although this holistic approach seems promising, scholars have only recently engaged in conceptual and empirical studies [2].

In addition, our study has implications for practitioners. The performance measurement tool can be used for several reasons. Social enterprises can use the developed performance measurement tool to deal with the tensions they are exposed to because of their hybrid character. First, we provide social enterprises with an internal self-assessment tool. Preferably, the assessment tool is completed by diverse employees of the organization. Their (diverse) opinions may give rise to an internal discussion about the non-financial performance of the organization. This may help social enterprises in preventing mission drift and safeguarding the balancing of social and financial goals in internal decision-making. Secondly, social enterprises can use the set of selected performance indicators as a guideline to report non-financial performance to external stakeholders. As such, they respond to the increasing demand for accountability, necessary to establish legitimacy.

Our paper also has some limitations, which have implications for future research. First, we conducted this study within Flemish social enterprises. It would be useful to replicate and generalize it to examine the validity of the developed tool on a larger, international scale. Furthermore, a promising trajectory for future research is to study the moderating effect of cultural differences in social enterprises between countries. As such, it would be interesting to examine the validity of this evaluation tool on a larger, international scale. Finally, performance management involves setting expectations for future achievements on the indicators which have been selected. The benefits of setting targets are that organizations have a focus and a clarity of organizational goals [20]. Future empirical studies could focus on setting targets, and could examine the impact of setting targets on organizational performance.

Acknowledgments: The study was made possible as a result of funding provided by The Policy Research Centre Work and Social Economy (Steunpunt WSE). The work was carried out with the active help of Flemish social enterprises. The support of these organizations is gratefully acknowledged. We would like to thank Tine Claeys and Linde Moonen. They were part of the research group and were involved in data gathering. We also would like to thank the two anonymous reviewers for their constructive feedback and helpful suggestions.

Author Contributions: Saskia Crucke and Adeliën Decramer conceptualized the framework of the study, collected the data and wrote the paper. Saskia Crucke analyzed the data.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix

Table A1. Overview of the Performance Domains of GRI, KLD, ISO26000 and DJSI [27,78,92,93].

	Global Reporting Initiative (GRI)	Kinder, Lydenberg, Domini (KLD)	ISO 26000	Dow Jones Sustainability Index (DJSI)
Economic	X	X Part of "Product"		X
Environmental	X	X	X	X
Community	X Part of "Category social": Human rights Product Responsibility Society	X Encompassing: Human Rights Product Community	X Encompassing: Human Rights Fair Operation Practices Consumer issues Community Involvement and Development	X Part of "Social dimension"
Human	X Part of "Category social": Labour practices and decent work	X Encompassing: Employee relations Diversity	X Labour Practices	X Part of "Social dimension"
Governance	X Part of "General standard disclosures"	X	X	X Part of "Economic dimension"

Table A2. Journals Screened in Phase 1 (Literature Review).

Journals screened on articles with "Social" and "Performance" in the title. Period: 1990–2013
Part 1: Academy of Management Journal Academy of Management Review Administrative Science Quarterly Business Ethics Quarterly Entrepreneurship: Theory & Practice Journal of Management Journal of Management Studies Organizational Science Organizational Studies Strategic Management Journal
Part 2: Journal of Business Ethics Social Enterprise Journal

References and Notes

1. Manetti, G. The role of blended value accounting in the evaluation of socio-economic impact of social enterprises. *VOLUNTAS: Int. J. Volunt. Nonprofit Organ.* **2014**, *25*, 443–464. [[CrossRef](#)]
2. Grieco, C.; Michelini, L.; Iasevoli, G. Measuring value creation in social enterprises: A cluster analysis of social impact assessment models. *Nonprofit Volunt. Sect. Q.* **2014**, *44*, 1173–1193. [[CrossRef](#)]
3. Arvidson, M.; Lyon, F. Social impact measurement and non-profit organisations: Compliance, resistance, and promotion. *VOLUNTAS: Int. J. Volunt. Nonprofit Organ.* **2014**, *25*, 869–886. [[CrossRef](#)]
4. Mair, J.; Marti, I. Social entrepreneurship research: A source of explanation, prediction, and delight. *J. World Bus.* **2006**, *41*, 36–44. [[CrossRef](#)]
5. Moss, T.W.; Short, J.C.; Payne, G.T.; Lumpkin, G.T. Dual identities in social ventures: An exploratory study. *Entrepren. Theor. Pract.* **2011**, *35*, 805–830. [[CrossRef](#)]
6. Santos, F.M. A positive theory of social entrepreneurship. *J. Bus. Ethics* **2012**, *111*, 335–351. [[CrossRef](#)]
7. Wilson, F.; Post, J.E. Business models for people, planet (& profits): Exploring the phenomena of social business, a market-based approach to social value creation. *Small Bus. Econ.* **2013**, *40*, 715–737.
8. Arena, M.; Azzone, G.; Bengo, I. Performance measurement for social enterprises. *VOLUNTAS: Int. J. Volunt. Nonprofit Organ.* **2015**, *26*, 649–672. [[CrossRef](#)]
9. Doherty, B.; Haugh, H.; Lyon, F. Social enterprises as hybrid organizations: A review and research agenda. *Int. J. Manag. Rev.* **2014**, *16*, 417–436. [[CrossRef](#)]
10. Smith, W.K.; Gonin, M.; Besharov, M.L. Managing social-business tensions: A review and research agenda for social enterprise. *Bus. Ethics Q.* **2013**, *23*, 407–442. [[CrossRef](#)]
11. Pache, A.-C.; Santos, F. Inside the hybrid organization: Selective coupling as response to conflicting institutional logics. *Acad. Manag. J.* **2013**, *56*, 972–1001. [[CrossRef](#)]
12. Battilana, J.; Metin, S.; Pache, A.-C.; Jacob, M. Harnessing productive tensions in hybrid organizations: The case of work integration social enterprises. *Acad. Manag. J.* **2014**, *34*, 81–100. [[CrossRef](#)]
13. Miller, T.L.; Grimes, M.G.; McMullen, J.S.; Vogus, T.J. Venturing for others with heart and head: How compassion encourages social entrepreneurship. *Acad. Manag. Rev.* **2012**, *37*, 616–640. [[CrossRef](#)]
14. Ebrahim, A.; Battilana, J.; Mair, J. The governance of social enterprises: Mission drift and accountability challenges in hybrid organizations. *Res. Organ. Behav.* **2014**, *34*, 81–100. [[CrossRef](#)]
15. Ramus, T.; Vaccaro, A. Stakeholders matter: How social enterprises address mission drift. *J. Bus. Ethics* **2014**, 1–16. [[CrossRef](#)]
16. Battilana, J.; Lee, M. Advancing research on hybrid organizing—insights from the study of social enterprises. *Acad. Manag. Ann.* **2014**, *8*, 397–441. [[CrossRef](#)]
17. Mair, J.; Mayer, J.; Lutz, E. Navigating institutional plurality: Organizational governance in hybrid organizations. *Organ. Stud.* **2015**, *36*, 713–739. [[CrossRef](#)]
18. Spear, R.; Cornforth, C.; Aiken, M. The governance challenges of social enterprises: Evidence from a uk empirical study. *Ann. Public Cooper. Econ.* **2009**, *80*, 247–273. [[CrossRef](#)]
19. Pache, A.-C.; Santos, F. When worlds collide: The internal dynamics of organizational responses to conflicting institutional demands. *Acad. Manag. Rev.* **2010**, *35*, 455–476. [[CrossRef](#)]
20. Boyne, G.A. Performance management: Does it work? In *Public Management and Performance*; Walker, R.M., Boyne, G.A., Brewer, G.A., Eds.; Cambridge University Press: Cambridge, UK, 2010; pp. 207–226.
21. Meadows, M.; Pike, M. Performance management for social enterprises. *Syst. Pract. Action Res.* **2010**, *23*, 127–141. [[CrossRef](#)]
22. Ebrahim, A.; Rangan, V.K. What impact? A framework for measuring the scale and scope of social performance. *Calif. Manag. Rev.* **2014**, *56*, 118–141. [[CrossRef](#)]
23. Nicholls, A. ‘We do good things, don’t we?’: ‘Blended value accounting’ in social entrepreneurship. *Account. Organ. Soc.* **2009**, *34*, 755–769. [[CrossRef](#)]
24. Bellucci, M.; Bagnoli, L.; Biggeri, M.; Rinaldi, V. Performance measurement in solidarity economy organizations: The case of fair trade shops in Italy. *Ann. Publ. Cooper. Econ.* **2012**, *83*, 25–59. [[CrossRef](#)]
25. Bagnoli, L.; Megali, C. Measuring performance in social enterprises. *Nonprofit Volunt. Sect. Q.* **2011**, *40*, 149–165. [[CrossRef](#)]

26. Van Loon, J.H.M.; Bonham, G.S.; Peterson, D.D.; Schallock, R.L.; Claes, C.; Decramer, A.E.M. The use of evidence-based outcomes in systems and organizations providing services and supports to persons with intellectual disability. *Eval. Program Plann.* **2013**, *36*, 80–87. [[CrossRef](#)] [[PubMed](#)]
27. GRI. Global reporting initiative. Available online: <https://www.globalreporting.org/resourcelibrary/GRIG4-Part1-Reporting-Principles-and-Standard-Disclosures.pdf> (accessed on 3 February 2016).
28. Dumay, J.; Guthrie, J.; Farneti, F. Gri sustainability reporting guidelines for public and third sector organizations. *Publ. Manag. Rev.* **2010**, *12*, 531–548. [[CrossRef](#)]
29. GRI. Ngo sector disclosures. Available online: <https://www.globalreporting.org/resourcelibrary/GRI-G4-NGO-Sector-Disclosures.pdf> (accessed on 3 February 2016).
30. Heras-Saizarbitoria, I.; Casadesús, M.; Marimón, F. The impact of iso 9001 standard and the efqm model: The view of the assessors. *Total Qual. Manag. Bus. Excel.* **2011**, *22*, 197–218. [[CrossRef](#)]
31. DeVellis, R.F. *Scale Development: Theory and Applications*; Sage: Thousand Oaks, CA, USA, 2003.
32. Hinkin, T.R. A review of scale development practices in the study of organizations. *J. Manag.* **1995**, *21*, 967–988. [[CrossRef](#)]
33. Hinkin, T.R. A brief tutorial on the development of measures for use in survey questionnaires. *Organ. Res. Meth.* **1998**, *1*, 104–121. [[CrossRef](#)]
34. Van Opstal, W.; Deraedt, E.; Gijssels, C. Monitoring profile shifts and differences among wises in flanders. *Soc. Enterp. J.* **2009**, *5*, 229–258. [[CrossRef](#)]
35. Defourny, J.; Nyssens, M. Conceptions of social enterprise and social entrepreneurship in europe and the united states: Convergences and divergences. *J. Soc. Entrepreneurship* **2010**, *1*, 32–53. [[CrossRef](#)]
36. Defourny, J.; Nyssens, M. Social enterprise in europe: Recent trends and developments. *Soc. Enterp. J.* **2008**, *4*, 202–228. [[CrossRef](#)]
37. Wood, D.J. Measuring corporate social performance: A review. *Int. J. Manag. Rev.* **2010**, *12*, 50–84. [[CrossRef](#)]
38. Laplume, A.O.; Sonpar, K.; Litz, R.A. Stakeholder theory: Reviewing a theory that moves us. *J. Manag.* **2008**, *34*, 1152–1189. [[CrossRef](#)]
39. Consolandi, C.; Jaiswal-Dale, A.; Poggiani, E.; Vercelli, A. Global standards and ethical stock indexes: The case of the dow jones sustainability stoxx index. *J. Bus. Ethics* **2009**, *87*, 185–197. [[CrossRef](#)]
40. Robinson, M.; Kleffner, A.; Bertels, S. Signaling sustainability leadership: Empirical evidence of the value of djsi membership. *J. Bus. Ethics* **2011**, *101*, 493–505. [[CrossRef](#)]
41. Hahn, R.; Lulfs, R. Legitimizing negative aspects in gri-oriented sustainability reporting: A qualitative analysis of corporate disclosure strategies. *J. Bus. Ethics* **2014**, *123*, 401–420. [[CrossRef](#)]
42. Levy, D.L.; Szejnwald Brown, H.; de Jong, M. The contested politics of corporate governance: The case of the global reporting initiative. *Bus. Soc.* **2010**, *49*, 88–115. [[CrossRef](#)]
43. Helms, W.S.; Oliver, C.; Webb, K. Antecedents of settlement on a new institutional practice: Negotiation of the iso 26000 standard on social responsibility. *Acad. Manag. J.* **2012**, *55*, 1120–1145. [[CrossRef](#)]
44. Balzarova, M.A.; Castka, P. Stakeholders' influence and contribution to social standards development: The case of multiple stakeholder approach to iso 26000 development. *J. Bus. Ethics* **2012**, *111*, 265–279. [[CrossRef](#)]
45. Wiklund, J.; Shepherd, D. Entrepreneurial orientation and small business performance: A configurational approach. *J. Bus. Venturing* **2005**, *20*, 71–91. [[CrossRef](#)]
46. Chang, H.-T.; Chi, N.-W. Human resource managers' role consistency and hr performance indicators: The moderating effect of interpersonal trust in taiwan. *Int. J. Hum. Resour. Manag.* **2007**, *18*, 665–683. [[CrossRef](#)]
47. Andersson, L.; Jackson, S.E.; Russell, S.V. Greening organizational behavior: An introduction to the special issue. *J. Organ. Behav.* **2013**, *34*, 151–155. [[CrossRef](#)]
48. Niehm, L.S.; Swinney, J.; Miller, N.J. Community social responsibility and its consequences for family business performance. *J. Small Bus. Manag.* **2008**, *46*, 331–350. [[CrossRef](#)]
49. Cornforth, C. Nonprofit governance research: The need for innovative perspectives and approaches. In *Nonprofit Governance, Innovative Perspectives and Approaches*; Cornforth, C., Brown, W.A., Eds.; Routledge: Abingdon, UK, 2014; pp. 1–14.
50. Hambrick, D.C.; Werder, A.; Zajac, E.J. New directions in corporate governance research. *Organ. Sci.* **2008**, *19*, 381–385. [[CrossRef](#)]

51. Daily, C.M.; Dalton, D.R.; Cannella, A.A. Corporate governance: Decades of dialogue and data. *The Acad. Manag. Rev.* **2003**, *28*, 371–382.
52. Chan, M.C.; Watson, J.; Woodliff, D. Corporate governance quality and csr disclosures. *J. Bus. Ethics* **2014**, *125*, 59–73. [[CrossRef](#)]
53. Huang, C.J. Corporate governance, corporate social responsibility and corporate performance. *J. Manag. Organ.* **2010**, *16*, 641–655. [[CrossRef](#)]
54. Arora, P.; Dharwadkar, R. Corporate governance and corporate social responsibility (csr): The moderating roles of attainment discrepancy and organization slack. *Corp. Govern. Int. Rev.* **2011**, *19*, 136–152. [[CrossRef](#)]
55. Cowton, C.J.; Downs, Y. Use of focus groups in business ethics research: Potential, problems and paths to progress. *Bus. Ethics Eur. Rev.* **2015**, *24*, S54–S66. [[CrossRef](#)]
56. Bruggen, E.; Willems, P. A critical comparison of offline focus groups, online focus groups and e-delphi. *Int. J. Market Res.* **2009**, *51*, 363–381. [[CrossRef](#)]
57. Caffey, R.H.; Kazmierczak, R.F.; Avault, J.W. *Developing Consensus Indicators of Sustainability for Southeastern United States Aquaculture*; Louisiana Agricultural Experiment Station, LSU Agricultural Center: Baton Rouge, LA; USA, 2001.
58. Landeta, J. Current validity of the delphi method in social sciences. *Technol. Forecast. Soc. Change* **2006**, *73*, 467–482. [[CrossRef](#)]
59. Okoli, C.; Pawlowski, S.D. The delphi method as a research tool: An example, design considerations and applications. *Inf. Manag.* **2004**, *42*, 15–29. [[CrossRef](#)]
60. Schmidt, R.C. Managing delphi surveys using nonparametric statistical techniques*. *Decis. Sci.* **1997**, *28*, 763–774. [[CrossRef](#)]
61. Rowe, G.; Wright, G.; McColl, A. Judgment change during delphi-like procedures: The role of majority influence, expertise, and confidence. *Technol. Forecast. Soc. Change* **2005**, *72*, 377–399. [[CrossRef](#)]
62. Von Der Gracht, H.A. Consensus measurement in delphi studies: Review and implications for future quality assurance. *Technol. Forecast. Soc. Change* **2012**, *79*, 1525–1536. [[CrossRef](#)]
63. Worrell, J.L.; Di Gangi, P.M.; Bush, A.A. Exploring the use of the delphi method in accounting information systems research. *Int. J. Account. Inform. Syst.* **2013**, *14*, 193–208. [[CrossRef](#)]
64. Spear, R. Governance in democratic member-based organisations. *Ann. Publ. Cooper. Econ.* **2004**, *75*, 33–59. [[CrossRef](#)]
65. Janssens, W.; Wijnen, K.; De Pelsmacker, P.; Van Kenhove, P. *Marketing Research with SPSS*; Prentice Hall: London, UK, 2008.
66. Hair, J.F.; Black, W.C.; Babin, B.J.; Anderson, R.E. *Multivariate Data Analysis*; Pearson: Upper Saddle River NJ, USA, 2006.
67. Rosseel, Y. Lavaan: An r package for structural equation modeling. *J. Stat. Software* **2012**, *48*, 1–36. [[CrossRef](#)]
68. Brown, T. *Confirmatory Factor Analysis for Applied Research*; The Guilford Press: London, UK, 2006.
69. Hu, L.; Bentler, P.M. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Struct. Equ. Model. Multidiscip. J.* **1999**, *6*, 1–55. [[CrossRef](#)]
70. Helm, S.T.; Andersson, F.O. Beyond taxonomy. An empirical validation of social entrepreneurship in the nonprofit sector. *Nonprofit Manag. Leader.* **2010**, *20*, 259–276. [[CrossRef](#)]
71. Mishra, S.; Suar, D. Does corporate social responsibility influence firm performance of indian companies? *J. Bus. Ethics* **2010**, *95*, 571–601. [[CrossRef](#)]
72. Chen, J.; Patten, D.M.; Roberts, R. Corporate charitable contributions: A corporate social performance or legitimacy strategy? *J. Bus. Ethics* **2008**, *82*, 131–144. [[CrossRef](#)]
73. Rettab, B.; Brik, A.; Mellahi, K. A study of management perceptions of the impact of corporate social responsibility on organisational performance in emerging economies: The case of dubai. *J. Bus. Ethics* **2009**, *89*, 371–390. [[CrossRef](#)]
74. O'Connor, M.; Spangenberg, J.H. A methodology for csr reporting: Assuring a representative diversity of indicators across stakeholders, scales, sites and performance issues. *J. Clean. Prod.* **2008**, *16*, 1399–1415. [[CrossRef](#)]
75. Graafland, J.; Eijffinger, S.C.W.; SmidJohan, H. Benchmarking of corporate social responsibility: Methodological problems and robustness. *J. Bus. Ethics* **2004**, *53*, 137–152. [[CrossRef](#)]
76. CAF. Common assessment framework. Available online: http://www.eipa.eu/files/File/CAF/CAF_2013.pdf (accessed on 3 February 2016).

77. Calantone, R.J.; Cavusgil, S.T.; Zhao, Y. Learning orientation, firm innovation capability, and firm performance. *Ind. Market. Manag.* **2002**, *31*, 515–524. [[CrossRef](#)]
78. ISO26000. Available online: <https://www.iso.org/obp/ui/#iso:std:iso:26000:ed-1:v1:en> (accessed on 3 February 2016).
79. De la Cuesta-González, M.; Muñoz-Torres, M.; Fernández-Izquierdo, M.Á. Analysis of social performance in the spanish financial industry through public data. A proposal. *J. Bus. Ethics* **2006**, *69*, 289–304. [[CrossRef](#)]
80. Heslin, P.A.; Vandewalle, D.O.N.; Latham, G.P. Keen to help? Managers' implicit person theories and their subsequent employee coaching. *Person. Psychol.* **2006**, *59*, 871–902. [[CrossRef](#)]
81. Milkie, M.A.; Peltola, P. Playing all the roles: Gender and the work-family balancing act. *J. Marriage Fam.* **1999**, *61*, 476–490. [[CrossRef](#)]
82. Herman, R.D.; Renz, D.O. Doing things right: Effectiveness in local nonprofit organizations, a panel study. *Publ. Admin. Rev.* **2004**, *64*, 694–704. [[CrossRef](#)]
83. Jackson, D.K.; Holland, T.P. Measuring the effectiveness of nonprofit boards. *Nonprofit Volunt. Sect. Q.* **1998**, *27*, 159–182. [[CrossRef](#)]
84. Fredette, C.; Bradshaw, P. Social capital and nonprofit governance effectiveness. *Nonprofit Manag. Leadersh.* **2012**, *22*, 391–409. [[CrossRef](#)]
85. Li, J.; Hambrick, D.C. Factional groups: A new vantage on demographic faultlines, conflict, and disintegration in work teams. *Acad. Manag. J.* **2005**, *48*, 794–813. [[CrossRef](#)]
86. Gill, M.; Flynn, R.J.; Reissing, E. The governance self-assessment checklist: An instrument for assessing board effectiveness. *Nonprofit Manag. Leadersh.* **2005**, *15*, 271–294. [[CrossRef](#)]
87. Minichilli, A.; Zattoni, A.; Zona, F. Making boards effective: An empirical examination of board task performance. *Br. J. Manag.* **2009**, *20*, 55–74. [[CrossRef](#)]
88. Wright, B.E. Public service and motivation: Does mission matter? *Publ. Admin. Rev.* **2007**, *67*, 54–64. [[CrossRef](#)]
89. Hillman, A.J.; Cannella, J.A.A.; Paetzold, R.L. The resource dependence role of corporate directors: Strategic adaptation of board composition in response to environmental change. *J. Manag. Stud.* **2000**, *37*, 235–255. [[CrossRef](#)]
90. Haynes, K.T.; Hillman, A. The effect of board capital and ceo power on strategic change. *Strat. Manag. J.* **2010**, *31*, 1145–1163. [[CrossRef](#)]
91. Claes, C.; Van Hove, G.; van Loon, J.; Vandeveld, S.; Schalock, R.L. Quality of life measurement in the field of intellectual disabilities: Eight principles for assessing quality of life-related personal outcomes. *Soc. Indic. Res.* **2010**, *98*, 61–72. [[CrossRef](#)]
92. KLD. *Kld Ratings Data: Inclusive Social Rating Criteria*; KLD Research & Analytics, Inc.: Boston, Massachusetts, USA, 2003.
93. DJSI. *The Dow Jones Sustainability World Index Guide*; S&P Dow Jones Indices LLC: New York, USA, 2012.



© 2016 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons by Attribution (CC-BY) license (<http://creativecommons.org/licenses/by/4.0/>).