Do good and measure well! – Examining the validity of two positive social change measurements in South African social enterprises

Philipp Kruse
Faculty of Psychology, Technische Universität Dresden, Dresden, Germany

Eleanor Meda Chipeta
Wits Business School, University of the Witwatersrand Johannesburg, Johannesburg, South Africa, and

Robert Venter
School of Business Sciences, University of the Witwatersrand Johannesburg, Johannesburg, South Africa

Abstract

Purpose – The creation of positive social change (PSC) is considered the primary success criterion when evaluating social enterprise performance. However, despite a proliferation of PSC-measurements, their empirical validity and applicability in emerging economies remain largely unclear. The quantitative study examines the validity of the PSC-measurement approaches proposed by Bloom and Smith (2010; Bloom and Smith approach [BSA]) and Weaver (2020b; Weaver approach [WA]) in South Africa.

Design/methodology/approach – Investigating a representative sample of 347 social entrepreneurs from Gauteng and Limpopo provinces, the authors use questionnaire data to explore the factorial, convergent and discriminant validity of both PSC-measurement approaches. Statistically, this is done by applying factorial and correlation analyses.

Findings – The results yield acknowledgeable differences. BSA has a high factorial and convergent validity, while its discriminant validity remains doubtful. For WA, problems concerning factorial validity occur.

Research limitations/implications – Despite limited generalizability, the authors provide a first guideline for scholars regarding the empirical validity of BSA and WA outside the context of developed economies.

© Philipp Kruse, Eleanor Meda Chipeta and Robert Venter. Published by Emerald Publishing Limited. This article is published under the Creative Commons Attribution (CC BY 4.0) license. Anyone may reproduce, distribute, translate and create derivative works of this article (for both commercial and non-commercial purposes), subject to full attribution to the original publication and authors. The full terms of this license may be seen at http://creativecommons.org/licenses/by/4.0/legalcode

The authors are thankful to the Saxon State and University Library Dresden (SLUB) for financially supporting Open Access publication of this study.

Funding: The research project was funded by the National Research Fund in collaboration with the University of the Witwatersrand.

Conflict of interests: None.
Originality/value – The current study sheds light on the validity of two PSC-measurement approaches in an emerging economy context. This way, the authors contribute to the field by addressing the scarcity of empirical research and the restricted scope of developed economies regarding PSC-measurement.

Keywords Positive social change, Social entrepreneurship, Measurement, Validity, South Africa

Paper type Research paper

Introduction
Since the Irish economist Cantillon (1756) put forward the first scientific conceptualization of entrepreneurial behaviour as acquiring goods at a certain price and selling them at a higher price (1,756), entrepreneurial performance measures have been dominated by financial indicators like growth, profit and liquidity [see Murphy et al. (1996) and Satalkina and Steiner (2020) for comprehensive overviews]. However, in the face of growing challenges worldwide like social inequality or environmental pollution, researchers and practitioners also acknowledge the potential of enterprises to address these challenges and contribute their share to societal well-being through the creation of positive social change (PSC) (Stephan et al., 2016; Fink, 2019; Bailey and Lumpkin, 2023).

While a growing number of commercially minded enterprises apply corporate social responsibility measures, e.g. by supporting local communities or charities (Kechiche and Soparnot, 2012), these measures are largely considered instrumental. They are intended to improve the enterprises’ reputation, attract prosocial investors and customers and, ultimately, improve financial performance (Aguinis and Glavas (2012) and Barnett et al. (2020) for comprehensive reviews). In contrast, social enterprises (SEs) consider the creation of PSC their primary mission (Kruse and Rosing, early online). SEs apply innovative entrepreneurial methods to address social challenges either as a for-profit SE (Kruse et al., 2021) or in a non-profit environment (Borzaga and Santucci, 2003). Thus, PSC-creation is the main success criterion of SEs, affects strategic decisions and the attraction of investors or donors (Dacin et al., 2010; Angelucci et al., 2023). Consequently, the measurement of PSC is a pivotal task for SEs. However, despite a proliferation of suggestions on how to measure SEs’ PSC (Grieco et al., 2015), research remains fragmented, hard to overlook and dominated by conceptual debates [see Rawhouser et al. (2019) for an overview].

Recently, Rawhouser et al. (2019) proposed a taxonomy of PSC-measurements distinguishing between (among others) outcomes for targeted beneficiaries, e.g. poverty reduction or job creation (Battilana et al., 2015; Tobias et al., 2013) and activities, e.g. the provision of public goods (Bai, 2013). The latter follows the logic that these activities will result in reaching the desired PSC. Furthermore, in addition to objective indicators, subjectively assessing enterprise performance from the entrepreneurs’ perspective is key to understanding the multi-facetted nature of entrepreneurial performance (Wach et al., 2016). This way, PSC-measures built on social entrepreneurs’ perceptions of enterprise success emerged. While the approach by Bloom and Smith (2010) encompasses scaling, i.e. PSC-growth with a clear outcome orientation, Weaver (2020b) focuses on the provision of different PSC-related activities by SEs. Whereas both approaches have great potential in PSC-measurement for SEs, two major shortcomings exist:

(1) Firstly, both approaches offer a possibility to quantify their understanding of PSC yet their empirical validity remains untested and thus unclear. On the one hand, Bloom and Smith (2010) adopted a global item approach and used their items to assess the criterion validity of proposed PSC antecedents. Thus, no comprehensive validation for these items took place (Diamantopoulos and Winklhofer, 2001). On the other hand, Weaver (2020b) remarks that her contribution rather lies in
clustering different PSC-related activities of SEs and a descriptive overlook regarding their provision. Consequently, the adaptation of Weaver’s findings as items and their validity examination are pending.

Secondly, both proposed measures were developed and applied in the USA. This mirrors the dominance of samples from Western and developed economies in behavioural sciences and business (Henrich et al., 2010; Arnett, 2008). However, SE-activity is particularly important in emerging economies like South Africa due to its great potential to alleviate social problems (Dupuy et al., 2016). Surprisingly, to date, research remains rather silent on how suitable the approaches by Bloom and Smith and Weaver are to measure PSC in emerging economies despite obvious differences to the highly developed USA economy.

Our research addresses these shortcomings by empirically examining the validity of both approaches in South Africa using a representative sample of 347 social entrepreneurs from Gauteng and Limpopo provinces. After outlining the main features of SEs, conceptualizing PSC and reviewing literature on PSC-measurement in the SE-context, we present the approaches by Bloom and Smith and Weaver in more detail. Subsequently, we empirically explore their factorial and construct validity using factor and correlation analyses and conclude by offering recommendations for future research, outlining implications for practitioners and society and reflecting our work’s limitations.

Theoretical background

Social entrepreneurship and positive social change

When introducing SEs as a new entrepreneurial form to a broader scientific community, Young (1983) entitled his book “If not for profit, for what?” implying that SEs are not primarily focused on financial but social value creation. Despite notable progress in the exploration of (nascent) social entrepreneurs’ characteristics and intentions (Kruse et al., 2021) and their personality and values (Tan et al., 2005; De Bernardi et al., 2023; Kruse et al., 2019), the answer to the question asked by young still remains vague. Broadly speaking, social entrepreneurs target the fulfilment of a social mission. This can be done in two ways. Firstly, SEs can be for-profit, i.e. they generate their own income based on an elaborated business model (Austin et al., 2006; Kruse et al., 2021). Secondly, they can operate in the non-profit sector generating income from government grants or public and private donations (Borzaga and Santuari, 2003; Felicio et al., 2013). Despite acknowledgeable differences in financing and the business model of for-profit and non-profit SEs (Kruse et al., 2023), the central commonality lies in their innovativeness when addressing the social problem identified (Mair and Marti, 2006; Dees and Anderson, 2006; Phills et al., 2008; Bailey and Lumpkin, 2023).

While assessing SEs’ success in the financing, i.e. monetary gains in the for-profit sector or the amount of money received through donations in the non-profit sector, is easy, the primary objective of for-profit and non-profit SEs, their PSC, remains hard to define and quantify (Dacin et al., 2010; Rawhouser et al., 2019). This is surprising given that PSC is key to determining the degree to which SEs achieve their desired aims and are successful (Bhatt, 2022).

Recent years saw a proliferation of terms referring to PSC with more or less overlap, like social performance (Nicholls, 2008), social returns (Emerson, 2003) or social value (Santos, 2012). Despite acknowledgeable nuances (Rawhouser et al., 2019; Hertel et al., 2020), we take an integrative and broad perspective on PSC and follow Stephan et al. (2016), who define PSC as “the process of transforming patterns of thought, behaviour, social relationships, institutions and social structure to generate beneficial outcomes for individuals,
Measuring positive social change in the social enterprise context

Reviewing SE-literature yields a myriad of approaches to assessing PSC. Just to name a few examples, Bai (2013) studies PSC in terms of the number of public goods provided by Californian hospitals. Tobias et al. (2013) measure PSC in terms of poverty reduction in the Rwandan coffee industry, whereas Corner and Ho (2010) study the effects of fair trade charity for Tibetans conducted by a SE. Also, Battilana et al. (2015) focus on job creation and investigate the PSC of a work integration SE as the percentage of beneficiaries finding a permanent job. An increasingly popular measure was proposed by Hall et al. (2015). The social return on investment (SROI) is, similar to the traditional return on investment, a possibility to quantify one’s PSC by calculating a ratio of costs relative to monetized effects, i.e. benefits provided by the SE. Granting that this can only be a very small selection of the huge number of PSC-assessments in the SE-context (Grieco et al., 2015 for an overview), the approaches presented illustrate the innate differences in measuring PSC.

In a comprehensive review, Rawhouser et al. (2019) proposed a taxonomy of PSC-measures. Among others, they distinguish between a focus on outcomes of PSC and a focus on PSC-activities. On the one hand, outcome-based operationalization’s focus on measuring the distinct results like poverty reduction or job creation (Battilana et al., 2015; Tobias et al., 2013). On the other hand, activity-based operationalization’s suppose that the activities conducted by the SE will result in reaching the desired effects (Bai, 2013). Thus, outcome measures are more immediate as distinct results of PSC are evaluated, while the activity logic suggests a mediate relationship of activity provision with (usually a broader range of) PSC-results.

Regardless of their orientation on outcomes or activities, PSC-measurement approaches are largely driven by stakeholder, government or investor/donor expectations to “objectivize” PSC (Hall et al., 2015) and find the best place to invest money and effort. This may be useful; however, there is a substantial risk that these metrics do not match the criteria of social entrepreneurs themselves. As Wach et al. (2016) showed, the perception of entrepreneurial success is broad and involves objective outcome measures that are more immediate as well as subjective criteria. To avoid imposing success criteria, entrepreneurs’ perceived PSC-success should receive (at least) the same attention as other metrics. Otherwise, despite being considered successful by stakeholders, particularly young (social) entrepreneurs might end up quitting, as their personal needs are not met (Bates, 2005).

Realizing the need to give social entrepreneurs themselves their say when measuring PSC, over the years, different self-report PSC-scales emerged (Grieco et al., 2015). While their goal, PSC-assessment from social entrepreneurs’ point of view, unites these approaches, their quality considerably varies. We take the view that, based on a review of several PSC-self-report measures, the two approaches by Bloom and Smith (2010) and Weaver (2020b) are particularly promising. They build on a PSC-definition in line with pertinent literature, apply a comprehensive theoretical framework to derive their measurements and attract
scholarly attention as a notable citation count, especially for the Bloom and Smith approach (BSA) shows (Weerakoon, early online).

The BSA follows the logic by Dees et al. (2008) and considers scaling, i.e. an increase of the quality and quantity of PSC created, a suitable criterion to assess PSC-success. However, in contrast to “objective” measures, the social entrepreneur him-/herself judges PSC-scaling compared to other, similar SEs over a three-year period. The judgement is done based on four items (see Table 1). BSA considers PSC-scaling a unidimensional construct and, following Rawhouser et al. (2019), has a clear outcome focus. Developed in the USA context, the items were first applied to 591 USA non-profit SEs to examine the criterion validity of a model of PSC-scaling antecedents. While, over the years, BSA-items or modified versions emerged as popular among PSC-scholars and were applied also in a for-profit SE-context (Nardini et al., 2022; Bacq and Eddleston, 2018), these studies do not provide an indication of the items’ factorial or construct validity and are limited to the USA context. So far, to the best of the authors’ knowledge, the application of BSA in emerging economies is rare and the few exceptions use it rather as a case study framework than an empirical self-report scale (Desa and Koch (2014) or Bocken et al. (2016), for examples).

The Weaver approach (WA) covers the provision of “opportunities pertaining to the advancement of human well-being” (Weaver, 2020b; p. 4), i.e. the activities offered by a SE to its beneficiaries. Weaver builds on the Capability Approach pioneered by Nussbaum (2004). This approach, that is widely recognized as an evaluative standard for social justice and PSC-creation in social entrepreneurship (Ziegler, 2010; Yujuico, 2008), comprises 13 central activities SEs can provide. WA clusters these activities into four thematic categories, as can be seen in Table 1. These clusters include improving mental and physical health of beneficiaries (Health and Human Security), improving the socioeconomic status of beneficiaries (Social Mobility), empowering beneficiaries to play an active role in society and politics (Social, Political and Environmental Engagement) and fostering beneficiaries’ creativity and relationship with other people (Self-Expression and Social Relationships). Considering the PSC-definition by Stephan et al. (2016) yields several notable overlaps of WA’s activities and Stephan et al.’s elements of PSC. Furthermore, an examination of 115 executive leaders of USA for-profit and non-profit SEs by Weaver (2020b) provided valuable insights regarding the amount and share of the 13 activities offered by these enterprises. However, so far, despite providing clear descriptions of the activities, they have not been adapted to formulate distinct items. Instead, WA-application is limited to theoretical reflections on how SE-capabilities may change in the face of crises (Weaver, 2020a; Weaver and Blakey, 2022) or case studies, e.g. in Indonesia (Duncan-Horner et al., 2022). Consequently, no empirical examination of WA’s validity as a self-report PSC-measure exists.

As this summary of BSA and WA yields, both approaches’ validity has not been examined with the empirical rigour necessary to judge important properties like factorial or construct validity. What is more, the context of development and application of BSA and WA is very limited to the USA context, i.e. one of the highest developed economies on the planet. As the literature review by Cao and Shi (2021) summarizes, through an entrepreneurship lens, significant differences comparing developed economies and emerging economies exist. Firstly, in emerging economies, institutional voids are widespread. As institutional theory (Scott, 1995) outlines, institutions can be formal (rules and laws) and informal (culture and norms) and affect the probability of successful entrepreneurial activity. Institutional void suggests that important resources for a successful entrepreneurial activity, like a reliable regulatory framework and a good (digital) infrastructure are constrained. Secondly, resource scarcities, i.e. the shortage of access to
### Table 1.
Summary of the PSC-measurements by Bloom and Smith (2010) and Weaver (2020b) and WA-item adaptations for our study

<table>
<thead>
<tr>
<th>BSA</th>
<th>Original formulations by Weaver (2020b, pp. 25–26)</th>
<th>WA-item adaptations used in our study</th>
<th>Thematic categories/latent factors with description by Weaver (2020b, pp. 25–26)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. I have made significant progress in alleviating the problem my enterprise addresses</td>
<td>(1) Health. Providing opportunities that benefit health (including reproductive), nourishment and life expectancy</td>
<td>1. Providing opportunities that benefit health (including reproductive), nourishment and life expectancy</td>
<td>Health and human security SE activities to develop and improve the mental and physical health, life expectancy and security of human beings</td>
</tr>
<tr>
<td>2. I have scaled up my capabilities to address the problem</td>
<td>(2) Mental/emotional health. Offering opportunities that foster mental health and that advance an individual’s ability to connect with others on an emotional level</td>
<td>2. Offering opportunities that foster mental health and that advance an individual’s ability to connect with others on an emotional level</td>
<td></td>
</tr>
<tr>
<td>3. I have greatly expanded the number of individuals my enterprise serves</td>
<td>(3) Safety or abuse. Providing opportunities that enable individuals to protect themselves from violent assault or cope after directly or indirectly surviving violent assault</td>
<td>3. Providing opportunities that enable individuals to protect themselves from violent assault or cope after directly or indirectly surviving violent assault</td>
<td></td>
</tr>
<tr>
<td>4. I have substantially increased the geographic area my enterprise serves</td>
<td>(1) Education. Providing services or products that advance educational development, including opportunities that foster critical thinking, imagination and reasoning</td>
<td>4. Providing services or products that advance educational development, including opportunities that foster critical thinking, imagination and reasoning</td>
<td>Social mobility SE activities aim to advance the socioeconomic status of individuals, social groups or communities. They involve developing the skills and abilities that enable people to increase their human capital, wealth and life goals</td>
</tr>
</tbody>
</table>

(continued)
### Table 1. Social, political and environmental engagement

<table>
<thead>
<tr>
<th><strong>BSA</strong></th>
<th><strong>Original formulations by Weaver (2020b, pp. 25–26)</strong></th>
<th><strong>WA-item adaptations used in our study</strong></th>
<th><strong>Thematic categories/latent factors with description by Weaver (2020b, pp. 25–26)</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>(1) Social issues and inclusion.</strong> Providing opportunities that enable people to organize around social issues and/or providing spaces of inclusion.</td>
<td>discrimination and/or to mobilize against such issues. 10. Providing opportunities that foster human interaction with things of nature, such as plants, animals and the overall environment.</td>
<td><strong>Social, political and environmental engagement</strong> SE activities involve creating opportunities for people to actively participate in society and the social systems that influence their social, political and environmental well-being.</td>
<td></td>
</tr>
<tr>
<td><strong>(2) Discrimination issues.</strong> Providing opportunities that enable individuals to deal with issues related to discrimination and/or to mobilize against such issues.</td>
<td>11. Providing opportunities that help people to engage in their political system, informing them of their rights and/or striving to protect their rights.</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>(3) Interaction with nature.</strong> Providing opportunities that foster human interaction with things of nature, such as plants, animals and the overall environment.</td>
<td>12. Providing opportunities that enable people to express themselves in a diversity of ways, including through art, religion and politics.</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>(4) Political participation.</strong> Providing opportunities that help people to engage in their political system, informing them of their rights and/or striving to protect their rights.</td>
<td>13. Providing opportunities that foster social interaction and participation in recreational activities that make them laugh or play.</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>(1) Independent/creative expression.</strong> Providing opportunities that enable people to express themselves in a diversity of ways, including through art, religion and politics.</td>
<td>Self-expression and social relationships. SE activities focus on fostering human creativity, recreation and their relationships with other human beings.</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>(2) Recreation or entertainment.</strong> Providing opportunities that foster social interaction and participation in recreational activities that make them laugh or play.</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
important resources like finance or a skilled workforce, are more frequent in emerging economies. This limits the capacities of nascent enterprises and makes them more vulnerable to failure. Thirdly, due to structural gaps, high-quality entrepreneurial education and support by organizations, e.g. in mentoring programmes, are less pronounced in emerging economies. Consequently, the level of entrepreneurship expertise is usually lower compared to developed countries with elaborated support infrastructures. From an SE point of view, first studies also suggest differences regarding SEs’ features (Kruse, 2021) and their prospects to successfully create PSC (Stephan et al., 2015) comparing SEs from developed and emerging economies.

To conclude, while BSA and WA show a high potential to add to the portfolio of PSC-measures by explicitly covering two different forms of PSC (outcomes and activities) and the social entrepreneurs’ subjective perceptions of PSC-creation, their factorial and convergent validity, as well as their suitability in an emerging economy context remain untested and thus unclear. Consequently, we formulate the following research question (RQ) to be examined in the current study:

RQ1. How valid are the approaches to measure PSC by Bloom and Smith (2010) and Weaver (2020b) regarding their factorial and construct validity when applied on social entrepreneurs operating in an emerging economy?

Methods

Sample choice, sampling technique, data acquisition and sample composition

To answer our RQ, we decided to collect primary data among social entrepreneurs in South Africa due to three reasons. Firstly, South Africa is an emerging economy and, as postulated by Cao and Shi (2021), nascent (social) entrepreneurs in the country face several challenges. These include corruption (Luiz and Stewart, 2014), resource scarcity reflected by high levels of inequality, chronic unemployment and crime (Littlewood and Holt, 2018), as well as a shortage of entrepreneurial hubs or mentoring programmes (Sheriff and Muffatto, 2015). Despite these challenges, a vivid SE-landscape developed in South Africa. Secondly, through a research lens, exploring PSC-creation in this context is particularly valuable as it is scientifically underexplored. Thirdly, from the perspective of policymakers, reliable empirical insights on SEs’ PSC-creation success are essential to use evidence-based decision-making practices when distributing limited (financial) resources and monitoring investment effectiveness.

To draw our sample, we combined purposive and random sampling techniques (Patton, 2002) in a two-step procedure. In the first step, we purposefully selected Gauteng and Limpopo out of the nine Provinces in South Africa. These two Provinces were chosen due to two reasons. On the one hand, Gauteng and Limpopo are considered very dynamic in their economic and social development and show high levels of entrepreneurial activity (Mwikikagile, 2008; Chakuzira, 2019). On the other hand, in contrast to the other Provinces, entrepreneurial activity in Gauteng and Limpopo is systematically recorded with a “Living List” generated by Gordon Institute of Business Science in collaboration with the Bertha Centre for Social Innovation and Entrepreneurship. The whole list features appx. 700 SEs and only contains social entrepreneurs acting as chief executive officers of their ventures. This way, we secured that only active creators of PSC were eligible for our study. Secondly, after obtaining the complete “Living List”, a random sampling technique was used. Random sampling was used as it is the most efficient and effective technique to secure that our final sample was representative for the total population of social entrepreneurs in the two Provinces. This increases the practical value of our findings for local policymakers (cf.
section on “Implications for Practice and Society”). Random sampling was conducted using a random number generator (Rotz et al., 2001). All SEs included in the “Living List” were numbered (1–700). Then, the generator randomly selected 50% of the total number of SEs, i.e. a sample of 350 SEs. This sample size is similar to other empirical studies on PSC (Bloom and Smith, 2010; Weaver, 2020b).

For data acquisition, at first, an email system was used to reach potential respondents. However, due to a low response rate and incomplete responses, physical addresses supplied in the database were used to locate the respondents face-to-face. Data were collected using a personal, structured and fully-standardized interview with the social entrepreneurs based on a questionnaire crafted by the authors of this study. The interviews were conducted by a contracted research company between January 2021 and May 2021. Out of the 350 SEs sampled, three could not be located reducing our sample to 347.

Our sample composition was as follows: 72.6% of social entrepreneurs led non-profit enterprises. The majority of enterprises has been existing for 16–20 years (36.6%), followed by enterprises ranging between 11 and 15 years (26.2%), 6–10 years (12.1%), more than 21 years (18.7%) and less than 6 years of existence (6.3%). Most enterprises had 11–20 employees (37.5%), followed by less than 10 employees (32.6%) and more than 20 employees (30.0%). The majority of social entrepreneurs’ age ranged between 36 and 45 years (37.2%), followed by social entrepreneurs aged between 20 and 35 years (31.4%) and social entrepreneurs older than 46 years (31.4%).

Validity examination
In line with our RQ, we investigated factorial and construct validity of BSA and WA. Factorial validity refers to the empirical data fit to a theoretically proposed factor structure. We examined factorial validity using two different methods of factor analysis (see section “Statistical Analyses” for more details). Construct validity, referred to as the extent to which theoretical constructs are well-operationalized and covered by the respective items (Bryant, 2000), has two elements, convergent and discriminant validity (Campbell and Fiske, 1959). Convergent validity signifies the overlap with already existing scales measuring similar constructs. Discriminant validity refers to an overlap with dissimilar constructs. Reviewing literature on suitable scales, the Subjective Entrepreneurial Success–Importance Scale (SES-IS) by Wach et al. (2016) emerged as particularly suitable for our purposes due to three reasons. Firstly, all in all, SES-IS is considerably short yet, consists of different sub-scales encompassing constructs similar and dissimilar to PSC. Thus, convergent and divergent validity can be examined with one concise scale keeping the questionnaire’s item number low. This decreases the risk for dropouts or incomplete responses. Secondly, SES-IS was crafted and validated in a multiple-step procedure using qualitative and quantitative methods. Thus, the scale possesses a high psychometric quality. Thirdly, during the validation process, SES-IS demonstrated its cross-cultural invariance, that is particularly important for instruments applied in different national and economic settings. Despite SES-IS not being used in South Africa so far, one can expect a high intercultural robustness.

In our study, we investigated convergent validity as follows: Firstly, we used two items from the SES-IS sub-scale “community impact” signifying the extent to which the entrepreneur considers his/her enterprise successful in creating social and community value. The sub-scale (α = 0.63) was rated on a seven-point Likert scale ranging from 1 (not successful) to 7 (very successful). An example item was “Please indicate the overall success of your enterprise regarding its social contribution”. Secondly, we used BSA to examine convergent validity of WA and vice versa. While in the study by Bloom and Smith (2010), distinct items are suggested that could be used in our study, the clustered 13 activities
proposed in WA had to be formulated as items to be included. Doing so, we used a three-step procedure. Firstly, the first and second authors of this study independently formulated items based on the description by Weaver (2020b). Secondly, both authors compared their formulations, discussed deviations and solved them to mutual satisfaction. Thirdly, the revised items were presented to an external expert in the field of PSC for her to comment on. Including these remarks resulted in the final items that can be found in Table 1 in the “WA-item adaptations used in our study” column.

Discriminant validity was examined with three SES-IS sub-scales rated on a seven-point Likert scale ranging from 1 (not successful) to 7 (very successful). For each sub-scale, two items were used: financial performance ($\alpha = 0.62$) refers to the financial sustainability of the enterprise. The items were adapted to fit both business models (for-profit and non-profit) of the enterprises investigated. An example item was “Please indicate the overall success of your enterprise regarding ways to finance your enterprise’s activities and keep it sustainable”. Workplace relationships ($\alpha = 69$) indicate the extent to which the working climate in the enterprise is perceived as positive (e.g. “Please indicate the overall success of your enterprise regarding employee satisfaction”). Finally, personal fulfilment ($\alpha = 0.62$) encompasses the satisfaction of the entrepreneur with his/her career, e.g. “Please indicate the overall success of your enterprise regarding opportunities for your personal development”.

Statistical analyses
Testing factorial validity, we first conducted a factor analysis using IBM SPSS 27. Doing so, item communalities were calculated to quantify the factor loadings of each item on an assigned latent factor. For BSA, only one latent factor is proposed by Bloom and Smith (2010). In contrast, Weaver (2020b) identified four thematic categories. As each category features at least two activities, they were treated as latent factors in our analyses. Computing the factor analysis, we followed Costello and Osborne (2005) and Velicer and Fava (1998) and checked whether item communalities were $\geq 0.70$ and loading on one distinct latent factor. This way, we identified potential item cross-loadings. Internal consistency of all latent factors was investigated with Cronbach’s alpha, that should be higher than 0.65 to be acceptable (Hair et al., 2009). Secondly, we applied structural equation modelling (SEM) with SPSS Amos to calculate the single items’ factor loadings on the underlying latent factor (Byrne, 2010). The assessment of overall empirical fit was based on the most frequent SEM model-fit indicators and corresponding thresholds [see Hooper et al. (2008) for an overview]. Construct validity was examined using correlation analysis. Regarding convergent validity, BSA and WA should have significant yet medium-size correlations with the SES-IS subscale “community impact” and the other respective PSC-measurement approach ($r \leq 0.65$), whereas, for discriminant validity, correlations with the other SES-IS sub-scales should be insignificant or small (Campbell and Fiske, 1959).

Results
Factorial validity
As can be seen in Table 2, all item communalities of BSA range above 0.70, which is an indication of a good factorial validity (Costello and Osborne, 2005; Velicer and Fava, 1998). Also, Cronbach’s alpha ($\alpha = 0.82$) yields a high internal consistency (Hair et al., 2009). SEM-analysis showed factor loadings above or close to the threshold of 0.70. Only item two scored notably lower at 0.62. A good overall fit regarding comparative fit index ($CFI = 0.96$), Tucker–Lewis index ($TLI = 0.96$) and standardized root mean residual ($SRMR = 0.04$) can be remarked. In contrast, root mean square error of approximation ($RMSEA (90\%$
confidence interval (CI) = 0.17 (0.11–0.24]) and \( \chi^2 \)-value \( \chi^2 (2) = 22.69, p < 0.01 \) are not in line with established thresholds of good model fit (Byrne, 2010; Hooper et al., 2008).

Examining WA, we find eight items with communalities \( \geq 0.70 \), three communalities ranging between 0.60 and 0.69 and two communalities below 0.60 and with notable cross-loadings (see Table 3). Furthermore, not all items fulfilling the criteria of high factorial validity (communality \( \geq 0.70 \)) load on the latent factor they are assigned to. For latent factor one (“Health and Human Security”), only item two has a high communality, whereas item one has high cross-loadings and item three loads with 0.70 on latent factor two (“Social Mobility”). For latent factor two, item five and item six have high communalities (0.81; 0.79) and item four and item seven score above 0.60. However, item four and item seven load on latent factor one. For latent factor three (“Social, Political, and Environmental Engagement”), communalities of items eight and nine fulfil the criteria of high factorial validity. Item 11 (0.69) is close to the threshold of 0.70; however, loading on latent factor one. The communality of item 10 is below 0.60. For latent factor four (“Self-Expression and Social Relationships”), both item communalities indicate a high factorial validity on the assigned latent factor. As Table 3 shows, the latent factors’ alphas range between \( \alpha = 0.66 \) (“Social Mobility”) and \( \alpha = 0.85 \) (“Health and Human Security”), indicating acceptable to high internal consistencies. Concerning SEM-results, we faced the problem that latent factor four could not be analysed at all, as the number of items assigned was insufficient to construct identified SEM-models (Byrne, 2010). For latent factor one, only factor loadings but no model fit indices could be calculated, as the number of items was too low. Regarding factor loadings, only two items scored above the threshold of 0.70, with two-factor loadings being close at 0.68. The other loadings remained notably below this threshold. The results of empirical model fit investigation for latent factor two [“Social Mobility”; \( CFI = 0.95; \ TLI = 0.94; \ RMSEA (90\% CI) = 0.12 (0.06–0.19); \ SRMR = 0.04; \chi^2 (2) = 12.11, p < 0.01 \)] indicate a reasonable fit while results for latent factor three [“Social, Political, and Environmental Engagement”; \( CFI = 0.99; \ TLI = 0.99; \ RMSEA (90\% CI) = 0.00 (0.00–0.05); \ SRMR = 0.01; \chi^2 (2) = 0.27 p > 0.05 \)] suggest a very good fit.

**Construct validity**

As can be seen in Table 4, BSA correlates significantly positive with all other constructs \( (p < 0.01) \). Regarding convergent validity, the correlation with community impact as measured by SES-IS is in a medium range \( (r = 0.49) \), indicating a notable yet not too big overlap with the construct (Campbell and Fiske, 1959). The same applies to the correlations with the WA-latent factor “Social, Political, and Environmental Engagement” \( (r = 0.50) \). The correlations with the other three latent factors are higher \( (0.56 \leq r \leq 0.64) \) yet still below

### Table 2.

Summary of item communalities and SEM factor loadings for the PSC-measurement approach proposed by Bloom and Smith (2010)

<table>
<thead>
<tr>
<th>Item</th>
<th>Communality</th>
<th>Factor loading in SEM</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. I have made significant progress in alleviating the problem my enterprise addresses</td>
<td>0.90</td>
<td>0.90</td>
</tr>
<tr>
<td>2. I have scaled up my capabilities to address the problem</td>
<td>0.71</td>
<td>0.62</td>
</tr>
<tr>
<td>3. I have greatly expanded the number of individuals my enterprise serves</td>
<td>0.84</td>
<td>0.75</td>
</tr>
<tr>
<td>4. I have substantially increased the geographic area my enterprise serves</td>
<td>0.78</td>
<td>0.68</td>
</tr>
</tbody>
</table>

**Notes:** Communalities after extraction of one factor are shown; SEM: structural equation model; internal consistency: \( \alpha = 0.82; N = 347; \) PSC: positive social change

**Source:** Authors’ own work
<table>
<thead>
<tr>
<th>Item</th>
<th>Item Description</th>
<th>Communalities</th>
<th>FL in SEM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Health and Human Security (LF 1)</td>
<td>1. Providing opportunities that benefit health (including reproductive), nourishment and life expectancy</td>
<td>0.49</td>
<td>0.46</td>
</tr>
<tr>
<td></td>
<td>2. Offering opportunities that foster mental health and that advance an individual’s ability to connect with others on an emotional level</td>
<td>0.70</td>
<td>0.30</td>
</tr>
<tr>
<td></td>
<td>3. Providing opportunities that enable individuals to protect themselves from violent assault or cope after directly or indirectly surviving violent assault</td>
<td>0.29</td>
<td>0.70</td>
</tr>
<tr>
<td>Social Mobility (LF 2)</td>
<td>4. Providing services or products that advance educational development, including opportunities that foster critical thinking, imagination and reasoning</td>
<td>0.65</td>
<td>-0.07</td>
</tr>
<tr>
<td></td>
<td>5. Providing opportunities that enable individuals to make plans towards reaching their life goals, as well as to critically reflect on them</td>
<td>0.13</td>
<td>0.81</td>
</tr>
<tr>
<td></td>
<td>6. Providing opportunities that help people obtain and own property, including houses, cars and other material goods</td>
<td>0.79</td>
<td>0.21</td>
</tr>
<tr>
<td></td>
<td>7. Providing opportunities that enable people to prepare and/or obtain employment</td>
<td>0.22</td>
<td>0.61</td>
</tr>
<tr>
<td>Social, Political and Environmental Engagement (LF 3)</td>
<td>8. Providing opportunities that enable people to organize around social issues and/or providing spaces of inclusion</td>
<td>0.25</td>
<td>0.37</td>
</tr>
<tr>
<td></td>
<td>9. Providing opportunities that enable individuals to deal with issues related to discrimination and/or to mobilize against such issues</td>
<td>0.09</td>
<td>0.20</td>
</tr>
<tr>
<td></td>
<td>10. Providing opportunities that foster human interaction with things of nature, such as plants, animals and the overall environment</td>
<td>0.33</td>
<td>0.24</td>
</tr>
<tr>
<td></td>
<td>11. Providing opportunities that help people to engage in their political system, informing them of their rights and/or striving to protect their rights</td>
<td>0.69</td>
<td>0.36</td>
</tr>
<tr>
<td>SE and SR (LF 4)</td>
<td>12. Providing opportunities that enable people to express themselves in a diversity of ways, including through art, religion and politics</td>
<td>0.37</td>
<td>0.29</td>
</tr>
<tr>
<td></td>
<td>13. Providing opportunities that foster social interaction and participation in recreational activities that make them laugh or play</td>
<td>0.05</td>
<td>0.17</td>
</tr>
</tbody>
</table>

**Notes:** Communalities after extraction of four factors and varimax rotation are shown; SE and SR: self-expression and social relationships; FL in SEM: factor loading in structural equation model; communalities and factor loadings ≥0.70 are printed in italic; internal consistencies: α_LF_1 = 0.85; α_LF_2 = 0.66; α_LF_3 = 0.67; α_LF_4 = 0.70; N = 347; PSC = positive social change; LF = latent factor, each latent factor corresponds to one thematic category identified by Weaver (2020b); †: for SE and SR no SEM factor loadings could be computed as the model was unidentified

**Source:** Authors’ own work
A correlation higher than 0.65 is typically considered an indication for an overlap casting doubt over the distinctiveness of constructs (Campbell and Fiske, 1959). Concerning discriminant validity, BSA-correlations with financial performance, workplace relationships and personal fulfillment range between 0.53 (workplace relationships) and 0.69 (financial performance). Particularly, the correlation with financial performance indicates a substantially high overlap and rather low discriminant validity.

Regarding WA’s latent factors, we find that all correlations displayed in Table 4 are positive and significant ($p < 0.01$). For convergent validity, latent factors’ inter-correlations range between 0.42 and 0.77. With several inter-correlations scoring above 0.65, casting doubt over the distinctiveness of these latent factors. All correlations with BSA, however, remain below 0.65 (0.50 \( \leq r \leq 0.64 \)). The same applies to all correlations with community impact (0.47 \( \leq r \leq 0.64 \)), suggesting a reasonable to good convergent validity. For discriminant validity, in addition to the three high inter-correlations of WA’s latent factors, we find only one correlation above 0.65 for the latent factors and financial performance, workplace relationships and personal fulfillment ($r_{Health and Human Security/Workplace Relationships} = 0.66$). All other correlations range between 0.42 and 0.61, indicating a reasonable to good discriminant validity.

**Discussion**

While the quantification of PSC becomes increasingly important, especially for SEs, progress in the field is hampered by an untested and thus unclear empirical validity of proposed approaches and a restricted scope as existing approaches are largely crafted and applied in developed economies only. Using a representative sample of 347 South African social entrepreneurs operating in Gauteng and Limpopo provinces, we explored factorial and construct validity of two proposed PSC-measurement approaches – the outcome-oriented approach by Bloom and Smith (2010) and the activity-oriented approach by Weaver (2020b) – in the context of an emerging economy. To do so, we applied factor analyses and correlational analysis.

Regarding our RQ, we find mixed results for factorial and construct validity:

Concerning factorial validity, BSA’s item communalities all exceed the threshold of 0.70. The same applies to SEM factor loadings, with only one item scoring notably below the threshold. All in all, these results indicate a high factorial validity (Costello and Osborne, 2005; Velicer and Fava, 1998). Also, alpha (0.82) yields a high internal consistency (Hair et al., 2009).

<table>
<thead>
<tr>
<th>Construct</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. BSA</td>
<td>–</td>
<td>0.58**</td>
<td>0.64**</td>
<td>0.50**</td>
<td>0.56**</td>
<td>0.69**</td>
<td>0.53**</td>
<td>0.55**</td>
<td>0.49**</td>
</tr>
<tr>
<td>2. Health and Security</td>
<td>–</td>
<td>–</td>
<td>0.77**</td>
<td>0.70**</td>
<td>0.62**</td>
<td>0.50**</td>
<td>0.66**</td>
<td>0.54**</td>
<td>0.64**</td>
</tr>
<tr>
<td>3. Social Mobility</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>0.67**</td>
<td>0.58**</td>
<td>0.52**</td>
<td>0.61**</td>
<td>0.48**</td>
<td>0.54**</td>
</tr>
<tr>
<td>4. SPE-Engagement</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>0.42**</td>
<td>0.47**</td>
<td>0.51**</td>
<td>0.42**</td>
<td>0.49**</td>
</tr>
<tr>
<td>5. SE and SR</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>0.45**</td>
<td>0.50**</td>
<td>0.46**</td>
<td>0.47**</td>
</tr>
<tr>
<td>6. Financial Performance</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>0.40**</td>
<td>0.43**</td>
<td>0.42**</td>
</tr>
<tr>
<td>7. Workplace Relationships</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>0.38**</td>
<td>0.50**</td>
</tr>
<tr>
<td>8. Personal Fulfillment</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>0.52**</td>
</tr>
<tr>
<td>9. Community Impact</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

**Notes**: BSA = Bloom and Smith approach; SPE-engagement = social, political and environmental engagement; SE and SR = self-expression and social relationships; PSC = positive social change; **$p < 0.01$; $N = 347$

**Source**: Authors’ own work

---

**Table 4.** Inter-correlations of the PSC-measurement approaches by Bloom and Smith (2010) and Weaver (2020b) with indicators of subjective entrepreneurial success by Wach et al. (2016)
SEM-results further indicate a good empirical fit despite a relatively high RMSEA-score and a significant $\chi^2$-score (Byrne, 2010; Hooper et al., 2008). However, considering that our sample is quite large and the high sample size sensitivity of $\chi^2$-scores (Bentler and Bonnet, 1980; Browne and Cudeck, 1989), overall, SEM-results show a good empirical fit. Consequently, the factorial validity of BSA can be considered high.

Examining the factorial validity of WA, we found that, of the 13 items used in our study, only eight item communalities exceeded the threshold of 0.70. Furthermore, two of these eight items loaded on a factor they were not assigned to based on Weaver’s thematic categories. While thematic categories three (Social, Political and Environmental Engagement) and four (Self-Expression and Social Relationships) emerge with two items over 0.70 each, thematic categories one (Health and Human Security) and two (Social Mobility) “keep” only one theoretically assigned item and “lose” one to the respective other scale. For all thematic categories, items with notable cross-loadings can be remarked. Despite reasonable to high internal consistencies of all four thematic categories ($66 \leq \alpha \leq 0.85$), these results cast doubt over WA’s factorial validity. Also, SEM-calculations yielded that only two items exceeded the threshold of 0.70, while another two came close at 0.68. The other factor loadings were notably lower or could not be calculated due to an insufficient number of items to reach an identified SEM-model (Byrne, 2010). Regarding empirical model fit, SEM-results for social mobility and social, political and environmental engagement are reasonable to very good yet no SEM-based examination of the other thematic categories was possible. To conclude, based on the items derived in our study, factorial validity of WA is weaker compared to BSA.

Examining construct validity, the results show that, for convergent validity, BSA has a significant yet not too high overlap with community impact as measured by SES-IS and all four of WA’s thematic categories. This yields a good convergent validity. WA displays some high inter-correlations of its thematic categories, with the highest reaching $r = 0.77$. However, the correlations of all thematic categories with BSA and community impact remain below 0.65, indicating a good convergent validity. Regarding discriminant validity, BSA-correlations with SES-IS subscales range between $r = 0.53$ (workplace relationships) and $r = 0.69$ (financial performance). Particularly regarding financial performance, this is more than one would expect of distinct constructs (Campbell and Fiske, 1959), indicating rather low discriminant validity. For WA, apart from one exception, all correlations of the four thematic categories with financial performance, workplace relationships and personal fulfilment remain in a medium range. However, as mentioned, some high inter-correlations amongst the thematic categories emerged, probably reflecting the problematic factorial validity. As a result, WA’s discriminant validity can be considered reasonable to good (Campbell and Fiske, 1959).

In sum, the current examination of BSA’s and WA’s factorial and construct validity suggests that BSA can be considered of a high factorial and convergent validity, yet, regarding discriminant validity slight problems occur. WA’s performance is weaker regarding factorial and construct validity. Particularly, the assignment of single items to the four thematic categories as latent factors seems empirically unsolid.

Implications for research
The current study has the potential to contribute to future research in three central ways:

1. Firstly, our work can be beneficial for hands-on empirical researchers looking for a suitable PSC-operationalization. As the results of the validity examination revealed notable differences comparing BSA and WA, the decision on which scale to choose should be driven by study goals and context. We recommend the application of
BSA in situations demanding a concise and economical assessment (as BSA consists of only four items), in which the degree of PSC is more important than its kind (as BSA has a clear outcome orientation) and that require the valid integration of single items to a higher-order factor (as BSA shows high factorial and convergent validity). Examples for such research projects include large panels that require short yet valid scales or quantitative studies targeting the application of advanced statistical methods like SEM (Kruse et al., 2023). Such methods demand high factorial validity of the scales used. WA can be useful in exploratory research. As WA depicts different kinds of PSC, it is well-suited for mapping PSC-activities in certain regions (provinces, countries, etc.). Moreover, applying WA-items in in-depth interviews could be helpful to identify (potentially) different strategies social entrepreneurs apply to create different kinds of PSC. Obviously, also the joint application of BSA and WA has great potential, as both perspectives (outcomes and activities) are covered. Combining BSA and WA results can be used to explore the question whether there are differences in the degree of PSC-creation in different activities (e.g. comparing health and social mobility).

(2) Secondly, the current study can contribute to improving the quality and robustness of empirical PSC-research. The proliferation of PSC-measurements that are theoretically sound yet not undergoing a rigorous validity examination (Grieco et al., 2015) could emerge as a major problem due to ambiguous results casting doubt on their empirical suitability. To illustrate, a recent meta-analysis by Kruse et al. (2021) uncovered that a fundamental share of heterogeneity in studies on social entrepreneurial intention originates from the application of different scales with questionable validities. We hope that our paper raises PSC-scholars’ awareness for the importance of applying valid PSC-measurements, as a good methodology is key to producing reliable, replicable and meaningful results (Kenny and Kashy, 1992).

(3) Thirdly, by investigating a sample of South African social entrepreneurs, our study overcomes the restricted scholarly scope on social entrepreneurs working in developed countries. It is known that, particularly in behavioural sciences, there is a disproportionate dominance of samples originating from Western, educated, industrialized, rich and democratic societies (Henrich et al. (2010) for an overview). PSC is no exception, as the majority of research stems from the USA (Rawhouser et al., 2019). However, as social entrepreneurs in non-Western contexts like South Africa create PSC under considerably different institutional circumstances (Cao and Shi, 2021), research here can be particularly insightful and should be intensified. For us, one particularly promising field of research features comparisons of PSC-creation by SEs operating in developed vs developing countries. This way, we could gain knowledge on differences, e.g. regarding the rule of law or cultural values and their impact on PSC. This, however, requires further validation of both approaches across other socio-economic conditions and the development of valid BSA- and WA-versions in different languages.

Implications for practice and society
In practice, the application of BSA and WA could be helpful to quantify PSC-creation yet avoid a simplistic, “monetarized” perspective. It is widely acknowledged that the successful accomplishment of PSC is harder to measure than economic performance indicators like revenue (Chipeta et al., 2022). This poses a challenge for social entrepreneurs, who usually
find it more difficult to convince investors and policymakers from the success of their enterprise (Dorado, 2006). While approaches like SROI try to tackle this problem by “objectivizing” PSC, we take the view that combining different stakeholders’ (i.e. subjective) perspectives on PSC is needed to account for the innate complexity in PSC-creation. Both BSA and WA can be valuable tools, as they allow to quantify these different perspectives. To illustrate, both scales can be filled in by social entrepreneurs and their target groups/beneficiaries. This way, potential investors and policymakers gain information on two separate perspectives on PSC-success. This is essential as PSC can only manifest if social entrepreneurial actions impact their beneficiaries’ lives (Stephan et al., 2016). Thus, we believe that complementing objective criteria of (financial) SE performance with different stakeholders’ subjective perspectives on PSC-performance using BSA and WA has the potential to result in a more nuanced, evidence-based and realistic assessment of SE performance. Ultimately, this could lead to more investments and political support, higher levels of social entrepreneurial activity, more effective alleviation of social problems and a fairer society.

Specifically, for policymakers in Limpopo and Gauteng, our findings can provide a valuable source of information. Due to the representativeness of our sample for the two Provinces, the outcomes of SE-activity in this region BSA and different activities undertaken by social entrepreneurs to achieve these outcomes WA are comprehensively examined. Building on these insights, local policymakers could find it easier to effectively distribute their (financial) resources and convince investors of the success of social SEs in the two Provinces. Moreover, designing tailor-made support and networking programmes in which SEs with similar approaches and activities can interact and support each other could be facilitated (cf. Perrini et al. (2010) on the importance of networking in SE).

Limitations

Our work has the following limitations:

Firstly, our study only examines the empirical validity of two selected PSC-measurement approaches. Yet, the total number of proposed measurements is notably higher (Grieco et al., 2015) and more diverse, going beyond approaches focusing on social entrepreneurs’ perception of PSC-creation.

Secondly, as WA did not feature distinct items in its original version, they were crafted in our study. Despite our efforts to be as close to the original work by Weaver (2020b) as possible (see Table 1), our items are just one possible adaption. This should be kept in mind when evaluating the generalizability of our findings.

Thirdly, while our study shed light on the validity of two existing self-report PSC-measures (BSA and WA) and identified several problems, crafting a new scale was beyond this paper’s scope. Thus, the development of a comprehensive self-report PSC-scale with a high convergent and discriminant validity in developed and emerging economies is still pending. Fourthly, we examined a representative sample of social entrepreneurs from two Provinces in South Africa. Yet, considering the large economic, ethnic and cultural diversity in the country (Ghosh, 2001), findings cannot be generalized. Furthermore, South Africa is just one of many developing countries with an emerging and vivid SE-landscape. Thus, our findings are not representative for all emerging economies.

Fifthly, for construct validation, the SES-IS by Wach et al. (2016) was used. As this scale only covers the perception of success from the entrepreneur’s perspective, future studies should extend our work by including rather “objective” and external indicators of PSC like SROI. These could be used to investigate the criterion validity of BSA and WA.
Sixthly, using cross-sectional instead of longitudinal data, like in our study, bears the risk of a lower internal validity (White and Arzi, 2005). To account for this shortcoming, we encourage scholars to investigate BSA’s and WA’s validity over time. This would allow the identification of validity changes in successive measurements or an examination of psychometric properties like re-test reliabilities.

Seventhly, examining the beneficiaries’ perception of PSC-creation was beyond the scope of our study. However, their perspective as the target audience of the SE is crucial (Hertel et al., 2020) and should be taken into account in future research.

Summary
In the light of growing societal expectations to contribute their share to addressing challenges like poverty or marginalization, enterprises increasingly commit to the creation of PSC. Particularly, SEs are dedicated to creating PSC and see this as their primary mission. Despite the importance of PSC as a success criterion of SEs, studies examining the validity of proposed measurements remain scarce. Furthermore, there is a limitation of scope, as PSC-measurements are almost exclusively crafted and applied in developed economies. Our study investigates the factorial and construct validity of two PSC-measurement approaches with a sample of 347 social entrepreneurs from the emerging economy of South Africa. The results of factor and correlational analyses yield notable validity differences comparing the approaches by Bloom and Smith (2010) and Weaver (2020b). Despite acknowledgeable limitations of the current study, our findings can serve as a guideline for scholars when deciding on a suitable PSC-measurement. We hope that, in the future, more effort is invested in the validity examination of proposed PSC-measurements and their application in emerging economies. This will contribute to a more reliable, more replicable and more diverse PSC research landscape.

References


**Corresponding author**
Philipp Kruse can be contacted at: philipp.kruse@tu-dresden.de