

DEVELOPING A CONCEPTUAL FRAMEWORK FOR COMPARING SOCIAL VALUE CREATION

ARNE KROEGER
CHRISTIANA WEBER
Leibniz University Hanover

The soaring popularity of business practices in the social sector has elicited numerous calls from academics and practitioners to adopt appropriate methodologies to quantify and compare social value creation. Contributing scholars consider it a great if not impossible challenge to compare social value creation of different, unrelated heterogenous interventions. We help bridge this research gap by developing a conceptual framework that allows us to compare the effectiveness of social interventions serving the different needs of different treatment groups in different socioeconomic and institutional contexts. We do so by bringing insights from both the literature on subjective well-being and the literature on organizational effectiveness theory into not-for-profit and social entrepreneurship research, as well as into the literature on program evaluation.

Measuring and comparing social value creation is an ongoing topic in society, politics, science, business, and economics. The importance of this topic is particularly great for not-for-profit organizations aiming to alleviate social problems such as hunger and poverty. Wondering whether they minister to their beneficiaries as effectively as possible, these organizations want to know not only “if they really are making progress and maximizing their potential positive social impact” but also “how they are performing relative to their industry peers” (Lingane & Olsen, 2004: 124). Such organizations often depend on material and immaterial support from third parties, including foundations, governmental institutions, corporations, and investors—sources that increasingly confront not-for-profit organizations with high expectations and demands regarding transparency and accountability (Campbell, 2002; Early, 2005; Grimes, 2010; Miller & Wesley, 2010; Nicholls, 2009; Polonsky & Grau, 2011). These sources seek to evaluate and compare a variety of social interventions to fund only the most effective ones (Buckmaster, 1999; Miller, Grimes, McMullen, & Vogus, 2012).

Measuring and comparing social value creation is presently being discussed in the literature on not-for-profit management, social entrepreneurship, and program evaluation (Bagnoli & Megali, 2009; Epstein & Klerman, 2012; Grimes, 2010; Holma & Kontinen, 2011; Polonsky & Grau, 2011; Sawhill & Williamson, 2001; White, 2013). Contributing

scholars consider it a great if not impossible challenge to compare social value creation of different, unrelated heterogenous interventions (e.g., Austin, Stevenson, & Wei-Skillern, 2006; Dacin, Dacin, & Matear, 2010; Emerson, 2003; Mair & Marti, 2006; Nicholls, 2009; Polonsky & Grau, 2011; Ryan & Lyne, 2008; Zahra, Gedajlovic, Neubaum, & Schulman, 2009). Indeed, our literature review revealed that current approaches neither clearly reflect social value creation nor specifically permit such comparisons.

To help close this research gap, we develop a conceptual framework for comparing the social value creation of different and unrelated interventions that serve different needs of different treatment groups in different socioeconomic and institutional contexts. To do so, we break with the dominant rationale by incorporating literature on (1) subjective well-being (SWB) and (2) organizational effectiveness theory into literature on (3) not-for-profit management and social entrepreneurship and (4) program evaluation.

More precisely, we use subjective satisfaction ratings to offer a uniform measurement unit that social interventions from different sectors can apply to gauge their performance and compare themselves with their industry peers. We then take the concept of *mean life satisfaction (LS) in regions and countries*, which conventionally indicates the living standard of regional or national economies, and combine it with insights from the not-for-profit literature. We conclude that social interventions primarily treat people

below regional or national levels of mean LS. We can thereby easily calculate the social need for different treatment groups in different regions or countries. Furthermore, we draw on organizational effectiveness theory—namely, the functional model (e.g., Cunningham, 1977: 468; Matthews, 2011: 84)—and adopt a basic understanding of value creation that enables us to posit a social intervention's effectiveness as a relative construct. Accordingly, we define the effectiveness of a social intervention as the degree to which an organization reduces a treatment group's social need. This degree can then be meaningfully compared with the degree of an entirely different social intervention that also reduces a treatment group's social need.

These combined elements from each of the literature streams within the various components of our framework bring us to new concepts of what "social" means in the not-for-profit context, of the social need in a region or country, and of the effectiveness of social interventions. These insights allow the comparison of unrelated heterogeneous social interventions. Our framework therefore parts from the prevailing view in the not-for-profit literature that such comparison is impossible. We also contribute to the not-for-profit and social entrepreneurship literature by introducing LS ratings—which conventionally indicate the living standard of regional or national economies—as a uniform measurement unit to assess the performance of interventions, regardless of the sector in which they occur. Moreover, our framework adds to the literature on program evaluation by providing a new method with which to contextualize program evaluation within the regional or national socioeconomic and institutional context for social interventions' operations. Finally, we contribute to the study of organizational effectiveness by reintroducing the functional model, which has lost attention in the literature, into the not-for-profit context.

This inquiry begins with an explanation of our understanding of a social intervention and the meaning of value creation in the social context. We then review the relevant literature, single out three central insufficiencies, and formulate three corresponding requirements of a framework for comparing social interventions that serve the different needs of different treatment groups in different socioeconomic and institutional contexts. We respond to these

requirements by drawing on the SWB literature and organizational effectiveness theory and by designing the framework. Finally, we discuss the contributions and outline theoretical and empirical implications of our proposed model.

SOCIAL INTERVENTIONS AND SOCIAL VALUE CREATION

Not-For-Profit Organizations and Social Interventions

Not-for-profit organizations conduct one or more social interventions. The ultimate purpose of social interventions is to create social value by improving the well-being of disadvantaged individuals (Martin & Osberg, 2007). One relatively new type of not-for-profit organization is the social enterprise, which searches for superior innovative solutions that can alleviate complex societal problems not sufficiently addressed by charities, governments, or commercial markets (e.g., Santos, 2012). This innovativeness often lies in a new combination of social and business practices. A well-known example of a social enterprise is a microfinance institute that runs a microfinance intervention intended to help women at the bottom of the pyramid break the vicious circle of poverty. The women receive microcredits and investment lessons that improve their well-being—specifically, their financial situations and investment skills. It is necessary to distinguish any improvement in the women's well-being from the intervention's pure outputs. The term *output* refers to the means: the microcredits and investment lessons that women receive to improve their financial situation and investment skills. This improvement, in turn, is the increase in women's well-being. We focus on the *ends*: those eventual changes in well-being. A discussion of the most effective and efficient means to initiate such changes is outside the scope of this article.

The development of a framework for comparing social value creation precedes a theoretically grounded understanding of value creation in the social context. Zahra et al. suggest clarifying the meaning of social value creation by integrating "insights from research on organizational effectiveness" (2009: 530). We follow this recommendation because organizational effectiveness affords an analytical framework that "represents a useful tool for critically evaluating

... the work of organizations" (Forbes, 1998: 183). This recommendation explicitly includes evaluation of the effectiveness of not-for-profit organizations (Herman, 1990).

Organizational Effectiveness Theory and Social Value Creation

The relevant literature on organizational effectiveness peaked between 1975 and 1985,¹ yielding "a potentially infinite number of effectiveness models" (Zammuto, 1984: 614). We stress "integrative, or multivariate, models [that] are generally more comprehensive and attempt to account for a greater proportion of the variance in effectiveness" (Steers, 1975: 547) than is the case with univariate effectiveness models, which use "only one measure of organizational success" (Steers, 1975: 546). In the relevant literature there have been several attempts to cluster and/or categorize different multivariate models of organizational effectiveness (e.g., Lewin & Minton, 1986; Quinn & Rohrbaugh, 1983). Here we follow the typology established by Yuchtman and Seashore (1967), two of the first scholars to discuss multivariate models. Their seminal publication has since been cited in highly rated academic contributions (e.g., Cameron, 1986; Connolly, Colon, & Deutsch, 1980; Cunningham, 1977; Lacey, Schmitz, & Swedlund, 2012; Lewin & Minton, 1986; Rojas, 2000; Sowa, Selden, & Sandfort, 2004; Zammuto, 1982) and identified three types of models: goal models, system-resource models, and functional models.

According to the goal model, effectiveness is the "degree to which an organization realizes its goals" (Etzioni, 1960: 262; see also Forbes, 1998). In other words, organizations aim to minimize the difference between intended and actual outcomes (Hannan & Freeman, 1977; Henri, 2004). In contrast, the system-resource model embeds the organization in its socioeconomic and institutional context. Organizations are effective to the degree that they assure their own viability (Forbes, 1998) by obtaining a "bargaining

position, as reflected in the ability of the organization ... to exploit its environment in the acquisition of scarce and value resources" (Yuchtman & Seashore, 1967: 898). In developing our framework for comparing social value creation, we draw on the third model, the functional model (e.g., Cunningham, 1977; Matthews, 2011), which contains elements of the other two and can therefore be regarded as a hybrid (Table 1).

This model places a goal orientation within a socioeconomic and institutional context and defines effectiveness as the *degree to which organizations' outputs "benefit the society to which they belong"* (Yuchtman & Seashore, 1967: 896; see also Connolly et al., 1980; Cunningham, 1977; Price, 1972). The functional model is therefore well suited to infer the effectiveness of social interventions that—as outlined above—benefit communities and society as a whole (Austin et al., 2006; Di Domenico, Haugh, & Tracey, 2010). Therefore, we consider the functional model to be the basis of our framework for comparing social value creation. However, the literature on organizational effectiveness has seldom explicitly mentioned the functional model because it overlaps with the goal and system-resource models. Instead, the functional model often has simply been subsumed by either the goal or system-resource model (Connolly et al., 1980; Price, 1972; Webb, 1974; Yuchtman & Seashore, 1967; Zammuto, 1982). Acknowledging the theoretically grounded basis that the functional model creates for understanding value creation in the social context, we now begin our review of existing approaches to measuring and comparing social value creation.

THE COMPLEXITY OF COMPARING SOCIAL VALUE CREATION

Review of Existing Approaches

Practitioners such as social analysts and managers of foundations and social enterprises have repeatedly tried to develop an appropriate framework for measuring and comparing social value creation (McLoughlin et al., 2009). The Trasi Database, created by the Foundation Center, contains a quite comprehensive collection of tools and resources for measurement and comparison (Foundation Center, 2012). In the academic community theoretical concepts for comparing social value creation are scarce.

¹ "From a total of more than 20 articles appearing on the topic of organizational effectiveness in the leading organizational science journals from 1975 to 1985 (and several hundred more before that time) [see Cameron & Whetten, 1983, for a review], only one article (Tsui, 1990) and one book (Denison, 1990) appeared during the next half decade" (Whetten & Cameron, 1994: 136).

TABLE 1
Similarities of Selected Effectiveness Models and Social Value Creation

Models and Characteristics of (Social) Value Creation	Characteristics of Social Value Creation	Goal Model	System-Resource Model	Functional Model
Posited outcomes	Degree to which a social intervention benefits society	Degree to which an organization minimizes the difference between intended and actual outcomes	Degree to which an organization is able to obtain an advantageous bargaining position to exploit its environment in acquiring scarce and valued resources	Degree to which organizations' outputs " benefit the society to which they belong" (Yuchtman & Seashore, 1967: 896; see also Cunningham, 1977, and Price, 1972)
Goal orientation	Benefiting society	Achieving organizational goals	Assuring an organization's viability	Benefiting society
Existence of socioeconomic and institutional contexts	Yes Social intervention is embedded into and influenced by a socioeconomic and institutional context	No Socioeconomic and institutional context is taken as a given	Yes Organization as part of a superordinate system whose subsystems help society run smoothly	Yes Organization as part of a superordinate system whose subsystems fulfill functions for society

Researchers (e.g., Bagnoli & Megali, 2011; Nicholls, 2008; Swanson & Di Zhang, 2010) have instead identified innovativeness, proactiveness, risk management, and other key dimensions as determining social value creation (Weerawardena & Mort, 2006). Perhaps the most sophisticated approaches are experimental and quasi-experimental research designs, such as randomized control trials (RCTs) or the difference-in-differences technique, as applied to program evaluation (e.g., Duflo & Kremer, 2008; Epstein & Klerman, 2012; Meyer, 1995; White, 2013). Although these approaches add substantial value to measuring the social value creation of social interventions, they "cannot be mechanically applied to other sites ... [because] institutions and communities in which programs are situated have their own histories and traditions" (Weiss, 1997: 513; see also Reinhardt, 2011). According to Woolcook, "Development professionals still lack a useable framework by which to engage in the vexing deliberations surrounding whether and when it is at least plausible to infer that a given impact result (positive or negative) 'there' is likely to obtain 'here'" (2013: 231).

Additionally, there is a vast body of social impact assessment (SIA) literature (e.g., Burdge & Vanclay, 1996; Esteves, Franks, & Vanclay, 2012; Finsterbusch, 1995; Lockie, 2001) that has arisen from the National Environmental Policy Act (NEPA) of 1969. Researchers in that stream understand SIA "as an umbrella or overarching framework that embodies the evaluation of all impacts on humans and on all the ways in which people and communities interact with their socio-cultural, economic and biophysical surroundings" (Vanclay, 2003: 5; see also Finsterbusch, 1985). Among other things, SIA puts an "emphasis on enhancing the lives of vulnerable and disadvantaged people" (Esteves et al., 2012: 40; see also Vanclay, 2003) and therefore requires consideration. Esteves et al. conclude that "there is strong consensus on what 'good' SIA practice looks like" (2012: 40). This consensus is reflected in the "Guidelines and Principles for Social Impact Assessment" (ICGP, 1994, 2003) and the 2003 "International Principles for Social Impact Assessment" (Esteves et al., 2012; Finsterbusch, 1995; Vanclay, 2003). However, these guidelines and the associated literature have a rather predictive (e.g., Finsterbusch, 1995) and descriptive nature and, thus, are of limited use for measuring and comparing the

value creation of social interventions (Esteves et al., 2012; Nicholls, 2009). According to Lockie's review of "the literature and practice of social impact assessment" (2001: 277), "comparative analysis offer[s] ... substantially less focus on precision—particularly in terms of mathematically reducing all impact variables to common denominators" (2001: 280). Lingane and Olsen add that "no framework has yet been articulated for quantifying the value of a company's impact on people and the environment" (2004: 116). In response, the authors elaborate on the social return on investment (SROI) approach (see also Emerson, 2003), which combines SIA practices and cost-benefit analysis.

SROI assesses the efficiency of an intervention. To calculate the SROI, analysts estimate the monetary equivalent of social value created and compare this equivalent to inputs used (Davis & Pett, 2002; Polonsky & Grau, 2011; Yates, 2009). The social value created is defined according to the intervention's customized theory of change. Hence, the logic of SROI offers valuable insights into an intervention's internal processes and drivers of social value. The method's downside is that different socioeconomic and institutional contexts and different treatment groups result in different theories of change for each evaluated intervention. Comparing two interventions on the basis of SROI requires similar interventions and becomes problematic with increasing heterogeneity in the socioeconomic and institutional contexts of the target groups whose social needs are served (Lingane & Olsen, 2004; Scholten, 2005). Two alternative approaches have been proposed, one by McLoughlin et al. (2009), who suggest a methodology for social impact measurement, and the other by Diochon and Anderson (2009), who develop a process typology for effectiveness in social enterprises. Both approaches do not allow comparison of heterogeneous interventions. Surprisingly, Diochon and Anderson's typology is not based on organizational effectiveness theory.

To summarize, although formalized measures might currently strengthen the credibility of frameworks developed by practitioners and academics, the many calls for appropriate methodologies to quantify and compare social value creation (e.g., Austin et al., 2006; Dacin et al., 2010; Emerson, 2003; Mair & Marti, 2006; Nicholls, 2009; Polonsky & Grau, 2011; Ryan & Lyne, 2008; Zahra et al., 2009) indicate that these current

approaches neither clearly reflect social value creation nor specifically permit comparisons.

Three Problems with Comparing Social Value Creation

A central conclusion of our literature review is that there seems to be consensus that comparability is hampered primarily by three characteristics of social value creation: (1) heterogeneous social interventions, (2) the social element, and (3) different socioeconomic and institutional contexts. First, there are myriad heterogeneous social interventions that operate in different areas of life to raise the well-being of different treatment groups (e.g., Austin et al., 2006; Colby, Stone, & Carttar, 2004; Dacin et al., 2010; Emerson, 2003; Nicholls, 2009; Paton, 2003; Polonsky & Grau, 2011; Reinhardt, 2011; Trelstad, 2008; Zahra, Rawhouser, Bhawe, Neubaum, & Hayton, 2008). Some researchers have considered it impossible to take a social intervention that equips hearing-impaired students in Europe with real-time, voice-to-text transcription and compare it to a microfinance intervention for raising the income of poor women in Bangladesh (e.g., Emerson, 2003; Polonsky & Grau, 2011). Analogously, it seems difficult to compare an intervention that gives nutrition paste to undernourished children in Kenya with one that provides newly released British prisoners with mentorship from rehabilitated ex-offenders to reduce the rate of recidivism. This heterogeneity of social interventions calls for a uniform social value construct that aligns the measurements of all of the different social interventions (Polonsky & Grau, 2011), and it requires an "essential" (Mair & Marti, 2006: 42) unit of measurement. Thus, the following is our first requirement:

- (1) A framework for comparing the social value creation of two independent interventions must build on a uniform social value construct and a unit of measurement that align the measuring of all the different social interventions.

Second, the "social" construct has not yet been defined in a manner capable of achieving a common understanding of what social value creation is. Therefore, the very intangibility of social value creation raises problems with comparing various forms of it (Austin et al., 2006; Choi & Majumdar, 2014; Dacin et al., 2010; Herman & Renz, 2008; Hervieux, Gedajlovic, &

Turcotte, 2010; Mair & Marti, 2006; Nicholls, 2008; Nicholls & Cho, 2008; Phills, Deiglmeier, & Miller, 2008; Polonsky & Grau, 2011; Tan, Williams, & Tan, 2005). In the literature on social entrepreneurship, social value created and commercial value created often come across as two extremes of a continuum (e.g., Chell, 2007; Dees, 1998). Scholars share the understanding that social value has been created if welfare improves, whereas commercial value creation has been achieved if profits ensue (Austin et al., 2006; Chell, 2007). However, commercial value creation might also improve welfare by creating jobs, for example (Austin et al., 2006; Hervieux et al., 2010), and social value creation might also improve an economic situation, as by generating earned income (e.g., Boschee, 2008; Chell, 2007). Therefore, the differentiation between social and commercial value created might hold only if one type of value created clearly dominates the other. In contrast, if social and commercial value creation are more balanced, discussions of the meaning of the concepts of social and commercial arise. For instance, objective judgment of whether a child day-care service or an e-learning platform for dyslexic students creates primarily social or mostly commercial value is problematic because such judgment depends on the individual evaluator's perception of when, at which point, and how social value is created (Austin et al., 2006).

To summarize, the current criteria for distinguishing between social and commercial value creation seem insufficient (Nicholls, 2008). Aside from this issue of differentiating between the two kinds of value creation, scholars (e.g., Dacin et al., 2010) have suggested typologies encompassing other sorts as well, such as cultural value creation. These additional types, too, might need to be distinguished from social and commercial value creation. However, the scope of this article is the comparison of social value creation. For that exercise evaluators require a common understanding of what social value creation means but not of the distinction between each other type of value creation. We therefore use the collective term *nonsocial* to subsume commercial value creation, cultural value creation, and all other potential types of value creation, and we confine our distinctions to social and nonsocial value creation. This proposal leads to our second requirement:

(2) A framework for comparing the social value creation of two independent interventions must distinguish between social and nonsocial value creation.

Third, the comparability of social value creation is also difficult where interventions have been undertaken in different socioeconomic and institutional contexts (e.g., Burdge & Vanclay, 1996; Kerlin, 2006; Mair & Marti, 2006; Reinhardt, 2011; Veenhoven, 2000; Zahra et al., 2009). People in different socioeconomic and institutional contexts have different living standards (Donaldson, 2001; United Nations, 2013). As a result, two interventions might increase their social value creation by the same amount, but the growth rate in the richer country will, by definition, be lower than the growth rate in the poorer country. Rates of improvement tend to be relatively high in low-number ranges, a fact that might exaggerate the effectiveness of a social intervention. For instance, a microfinance intervention that increases the monthly income of a group of rural women in Bangladesh on average from U.S. \$40 to U.S. \$45 at purchasing power parity (PPP)-adjusted exchange rates would achieve an income increase of 12.5%. If this intervention increased the monthly income of a group of rural women in Chile by the same amount of U.S. \$5, but from U.S. \$100 to U.S. \$105 at PPP-adjusted exchange rates, it would achieve only an income increase of 5%. It seems difficult to make sense of comparing improvements in social value created in different countries. For that reason social value creation must be contextualized within the environment where it has occurred. Thus, the following is our third requirement:

(3) A framework for comparing social value creation of two independent interventions must allow for the possibility of different socioeconomic and institutional conditions.

We continue developing our framework by responding to these three requirements.

A FRAMEWORK FOR COMPARING SOCIAL VALUE CREATION

Requirement 1: Uniform Social Value Construct and Measurement Unit

Social value created. Social interventions serve the different needs of particular treatment groups, such as raising the income of women at the

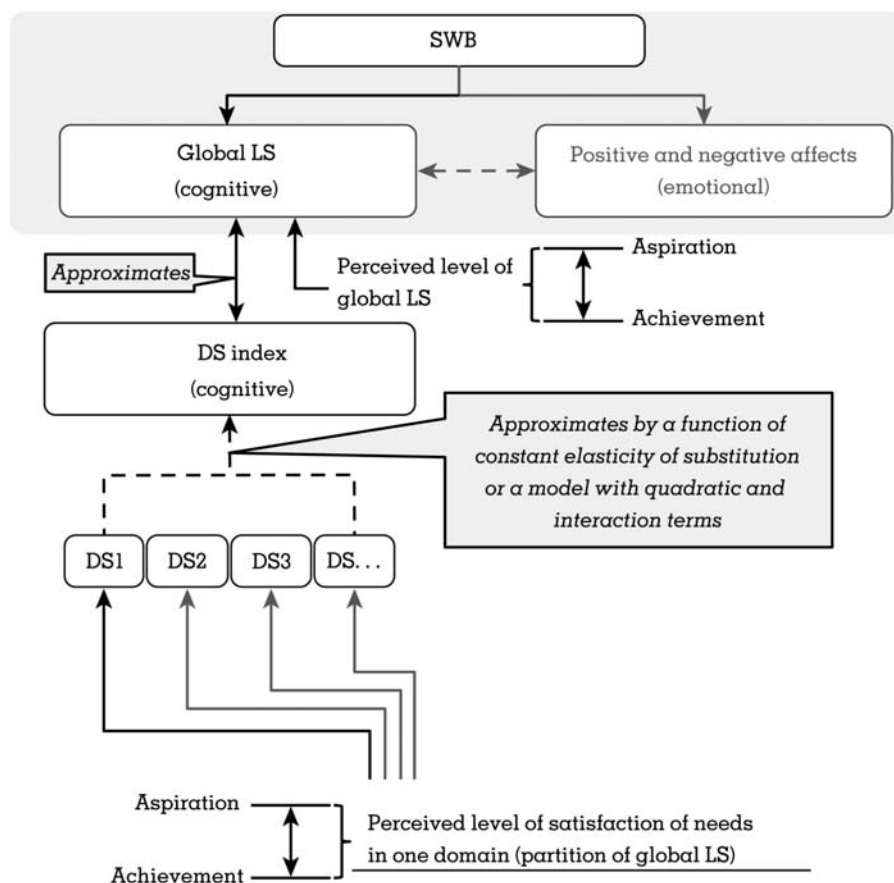
bottom of the pyramid by granting microcredits and developing investment skills, or bettering the health of slum inhabitants by improving sanitation. In line with the current literature, we understand social value created as the positive change, initiated by a social intervention, in the SWB of disadvantaged individuals (Clark, Rosenzweig, Long, & Olsen, 2004; Nicholls, 2009).

SWB. The SWB construct is “primarily concerned with the respondents’ own internal [perceptual] judgment of well-being, rather than what policymakers, academics, or others consider important” (Diener & Suh, 1997: 201). For the purpose of this article, we confine ourselves to the construct of SWB because it “is a well-researched and evaluated concept in terms of its validity” (Kroll & Delhey, 2013: 17). It has “high reliability” (Diener, Inglehart, & Tay, 2013: 498) and is well known in both economic theory and psychology (Layard, 2010; Veenhoven, 2000). Additionally, evaluations of SWB allow “respondents to freely weight different aspects. Thus, the scales can take people’s values and preferences into account as well as the outcomes of their choices” (Diener et al., 2013: 498). Furthermore, subjective measures elicit “very high response rates” (Kroll & Delhey, 2013: 16). Evidence for these very high response rates comes from the World Value Surveys (www.worldvaluessurvey.org/WVSONline.jsp). According to Diener et al. (2013), the World Value Surveys, which also assess SWB, have achieved a mean response rate of 98% in the last thirty years.

The SWB construct overlaps with the freedom-generative capabilities approach, which was developed by Nobel Laureate Amartya Sen. We therefore distinguish between these two approaches. In Sen’s (1985, 1999) approach opportunities are regarded as a key to the “well-being freedom” (1985: 201) construct, but we apply a conventional understanding of “being well” (Sen, 1985: 195) instead. We view opportunities as a means of raising SWB (see also Anand et al., 2009, and Veenhoven, 2009). Drawing on the relevant literature, we suggest the following structure of SWB, which clarifies not only our understanding but also the interrelationship between the different constructs of SWB (Figure 1).

Researchers have frequently highlighted the emotional and cognitive components of SWB (e.g., Campbell, Converse, & Rodgers, 1976; Diener, 1984; Erdogan, Bauer, Truxillo, & Mansfield, 2012; Schimmack, 2008). The emotional component

FIGURE 1
The Structure of SWB



Note: SWB = subjective well-being; LS = life satisfaction; DS = domain satisfaction; DS index = $f(DS_1, DS_2, \dots, DS_n)$.

of SWB is composed of both positive and negative affects. These affects indicate the degree to which an individual associates an experience with either pleasant or unpleasant feelings (Bradburn, 1969). The cognitive component of SWB concerns an individual's life satisfaction (LS; Andrews & Withey, 1974; Diener, 1984; Erdogan et al., 2012). Busseri and Sadava have discussed alternative interdependencies between LS and positive and negative affects. The authors "urge researchers interested in studying SWB [to examine LS, positive affects, and negative affects]... as separate components or LS as the ultimate criterion of interest" (2011: 302). In response to this recommendation, our framework for comparing social value creation accentuates the LS component of SWB. One reason for this choice is that individuals' reporting of LS is statistically more reliable over time than are their experiences of positive and negative feelings

(Diener et al., 2013). Moreover, the most widely used component of SWB is the "appreciation of life as-a-whole" (Kroll & Delhey, 2013: 17).

Global LS. In keeping with the understanding in the literature (e.g., Campbell et al., 1976; Carp & Carp, 1982; Layard, 2010), we define LS as the result of an individual's perceived discrepancy between his or her aspired and achieved levels of need. This definition emphasizes three facets of LS. First, LS is assessed at an individual level (Layard, 2010). Usually, evaluators ask individuals to rate their LS on a Likert scale (Anand et al., 2009; Beja & Yap, 2013; Finsterbusch, 1985; Georgopoulos & Mann, 1962; Lawless & Lucas, 2011; Layard, 2010; Price, 1972; Veenhoven, 2000). For instance, the Gallup World Poll asks people on which step of "an eleven-rung ladder where the bottom (0) represents the worst possible life... and the top (10) represents the best possible life" they feel they "personally stand at the present

time" (OECD, 2014: 134). Similarly, the World Value Surveys ask people to report their LS on a scale from 1 (*dissatisfied*) to 10 (*satisfied*), "taking all things together" (www.worldvaluessurvey.org/WVSONline.jsp). Individuals' LS can be aggregated at the level of either a group or society at large (Campbell et al., 1976). For example, data from the World Value Surveys make it possible to calculate the mean LS of representative samples from eighty-four countries. But it is important to distinguish between the mean LS of groups and the proportion of individuals within a group whose members report a certain level of LS. For instance, World Value Surveys data document that the mean LS of Bangladesh is 5.8, yet the proportion of Bangladeshi citizens who report an LS level of 6 is 9.3% (www.worldvaluessurvey.org/WVSONline.jsp, wave 4: v81). In this article we focus on the mean LS of groups, regions, and countries.

Second, LS is "a highly personal experience" (Campbell et al., 1976: 10). It is not socially constructed, for the discrepancy between a person's aspired and achieved levels of need is a result of the individual's own perceptions (e.g., Diener, 1984; Keyes, 1998). However, LS is determined by an individual's aspired level of need, which is considered to be socially constructed because it results from a process of social negotiation (Berger & Luckmann, 1967; Herman & Renz, 2008; Schuetz, 1967).

Third, LS refers to the needs of human beings. It does not include environmental damage. Although environmental value creation and social value creation are often treated as one construct (e.g., Burdge & Vanclay, 1996), we assume that environmental and social value creation do have strong interdependencies but follow different logics of value creation. We regard environmental aspects as determinants of an individual's satisfaction (e.g., air or water pollution causes a need for improved health).

Given these three characteristics of LS and the previously mentioned condition that the ultimate purpose of a social intervention is to enhance the LS of its treatment group(s), individuals' global LS may well constitute an appropriate basis for our framework for comparing social value creation. Although an enormous range of heterogeneous social interventions exists, this construct of global LS can subsume the different needs experienced by different treatment groups targeted by social interventions. A microfinance

intervention raises the LS of women at the bottom of the pyramid, a fair-trade intervention raises the LS of deprived farmers, and so on. The underlying assumption that one can sufficiently measure a social intervention's effectiveness by asking members of a treatment group about their LS has been proven "in hundreds of routine population surveys" (Layard, 2010: 534). However, a more direct link between a treatment group's particular needs and its LS ratings would allow one to attribute the outputs of a social intervention more precisely to the initiated LS of its treatment group (Cummins, 1996). A more direct link would also deepen the insights that politicians, researchers, and managers in the social sector have into how and why individuals judge their life situations in certain ways (Tay, Chan, & Diener, 2013) and would offer a "multifaceted picture" (Diener, 1984: 545).

DS. To produce a more direct link between the outputs of interventions and the improved LS of treatment groups, global LS can be segmented into several "partitions" (Rojas, 2006: 469), each of which is referred to as DS. Our understanding of domains refers to human needs, such as the need for mobility, safety, health, equality, community integration, and education (e.g., Carp & Carp, 1982; Diener, 1984; Hsieh, 2003; Stiglitz, Sen, & Fitoussi, 2009). Accordingly, DS is the satisfaction an individual feels with his or her perceived situation related to a particular need, examples of which include satisfaction with mobility and safety. For instance, the Australian Unity Personal Wellbeing Index contains the question, "On a scale of 0–10, how satisfied are you with *how safe you feel*?" (www.australianunity.com.au/about-us/wellbeing/auwbi; emphasis added).

The relationship between LS and DS is still under discussion in the literature. Bottom-up proponents (e.g., Campbell et al., 1976; Carp & Carp, 1982; Rojas, 2006) have posited that "life satisfaction judgments are based on an assessment of satisfaction in a relatively small number of life domains" (Schimmack, 2008: 2–3). Proponents of the top-down approach have assumed that changes in DS are determined solely by changes in LS (e.g., Diener, 1984; Erdogan et al., 2012; Schimmack, 2008). The bottom-up and top-down approaches merge in the bidirectional approach, whose supporters assume the existence of a mutual relationship between DS and LS (Gonzalez, Coenders, Saez, & Casas, 2010;

Mallard, Lance, & Michalos, 1997). Taking the bidirectional approach. Cummins, Eckersley, Pallant, Van Vugt, and Misajon (2003) obtained results suggesting that a DS index is not the same as LS but might meaningfully approximate it.

Hence, one can form a DS index by aggregating several DSs. The relationship between a DS index and its included DSs is expressed in the following formula: $DS\ index = f(DS_{1f}, DS_{2f}, \dots, DS_{nf})$. That is, a DS index is a function of a number of DSs and correlates with global LS. According to Rojas, the "enumeration and demarcation of the domains ... is arbitrary... Nevertheless, any partition must value parsimony—the number of domains must be manageable and domains should refer to clearly separable information—, meaning" (2006: 469). Cummins (1996) suggested that there should be at least seven domains to approximate LS appropriately. Beja and Yap have suggested that such an index is defined "by public policy or defined by people through, say, a survey" (2013: 631). Different interventions might address different DSs. If these different DSs are aggregated to one DS index no matter which DS is addressed, then all social interventions address this DS index.

Building on Cummins (1996), who conducted a meta-analysis of thirty-two studies on DS, one could design a fictive index consisting of seven DSs: satisfaction with education, financial situation, health, housing, community integration, equality, and safety. A microfinance intervention might seek to raise this DS index by addressing DS related to education, financial situation, and community integration. A social intervention that provides real-time, voice-to-text transcription to hearing-impaired students might raise the DS index by addressing hearing-impaired students' DS related to education and community integration. One advantage of a DS index that Diener et al. found is that "multi-item life satisfaction scales" (2013: 499) are more stable than a single-item LS. However, DS indices have prompted several discussions, particularly with regard to the aggregation of DS and the weighting of any particular DS.

With regard to the aggregation of DS, Rojas concluded that "most studies have assumed that the relationship between life satisfaction and domains-of-life satisfaction is additive" (2006: 470). However, he found both theoretical and empirical evidence that the additive

calculation is inferior to a function he calls "constant elasticity of substitution" (2006: 467), which takes into account synergies and substitution effects among different DSs. Another point is that Rojas's function includes decreasing marginal returns of LS in the presence of higher levels of DS. These effects also inform a model using quadratic and interaction terms (proposed by Gonzalez et al., 2010). We regard synergies, substitution effects, and decreasing marginal returns as necessary conditions of a DS index, and we therefore suggest choosing either a function of constant elasticity of substitution or a function with quadratic and interaction terms.

As for weighting DS, Locke (1969) documented that it is not necessary to account for weightings when aggregating DSs, as in a DS index. He supplied empirical evidence that the perceived level of DS—a result of the perceived discrepancy between aspired and achieved levels of need in one area of life—tends to be higher for domains of high importance than it is for domains of low importance. Locke's "results indicate that satisfaction with some object or situation is a function not only of the amount of discrepancy between percept and value, but of the importance of that value to the individual" (1969: 330). Wu (2008) has given empirical evidence of this view that weightings are dispensable. However, some researchers represented in the well-being literature (e.g., Hsieh, 2003) have included weightings in their formulas for aggregating DS. We follow Locke's (1969) argument and dispense with weighting of DSs.

In summary, LS and, more specifically, a DS index can function as social value constructs applicable to all social interventions intended to increase the satisfaction of a treatment group. Accordingly, the measurement of these constructs occurs on a Likert scale and can be applied by social interventions serving the different needs of their particular treatment groups. Theoretically, a DS index is a more precise and robust construct than global LS. In terms of practicality, LS might be the less complex construct with more comprehensive data sets already in place, such as those from the Gallup World Poll and the World Value Surveys. For these reasons we build our framework on the DS index but cite real LS examples for lack of DS index data.

Requirement 2: Social Element

Our second requirement for comparing the social value creation of two independent interventions centers on the meaning of social value creation. Developing a framework for comparing social value creation requires guidance on what distinguishes primarily social value creation from primarily nonsocial value creation. As outlined in the section on organizational effectiveness theory, the functional model suggests that social interventions create value if they benefit society through their output, for "from the point of view of the organization itself [it] is identical with, or corresponds to, effectiveness viewed from the vantage point of some other entity, such as a member, or owner, or the community, or the total society" (Yuchtman & Seashore, 1967: 895). However, social interventions usually do not treat every individual in a society. By adding up the current DS indices of a representative sample of individuals treated and not treated by the focal intervention and by dividing the resulting sum by the number of individuals in this sample, one can calculate the society's mean DS index (hereafter termed the *DS index threshold*).

In this article we follow a pragmatic approach, understanding the term society to mean the region or country in which the social intervention operates. A wide range of data on national and regional mean LS are already available, so no additional assessment of newly generated data are required to apply a national or regional threshold. National mean LS data come from the European values Study (www.europeanvaluesstudy.eu), the World Value Surveys (www.worldvaluessurvey.org/WVSONline.jsp), the OECD Better Life Index (<http://www.oecd.betterlifeindex.org/>), the UK's National Well-being Measures (www.ons.gov.uk/ons/datasets-and-tables/index.html), and Bhutan's Gross National Happiness Index (Ura, Alkire, Zangmo, & Wangdi, 2012), among others; regional mean LS data, for instance, come from the New Zealand General Social Survey (http://www.stats.govt.nz/browse_for_stats/people_and_communities/Well-being.aspx) and the Gallup Healthways Well-Being Index (<http://www.well-beingindex.com>). The choice between regional or national satisfaction thresholds depends mainly on data availability—namely, on whether national or regional satisfaction thresholds have already

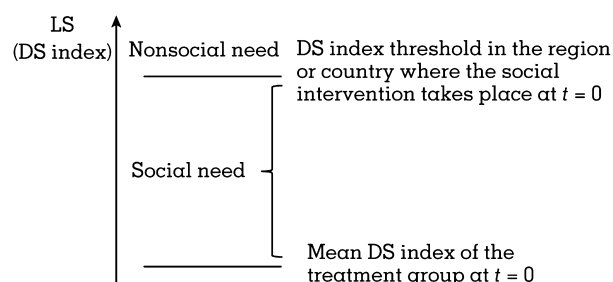
been evaluated. The DS index threshold thus represents the absolute (as opposed to relative) satisfaction level of the studied region or country. Individuals whose DS index falls below this threshold are considered disadvantaged relative to individuals whose DS index lies above it. One possible reason for the disparity is that it is more difficult for disadvantaged people to gain access to resources than it is for advantaged people (Reed, Jones, & Irvine, 2005). Alternatively, the DS disparity between these two groups might exist because the disadvantaged people have physical handicaps. Blind people, for instance, might perceive low satisfaction with their community integration, a response that lowers their DS index below the mean DS index of the region or country. A disadvantaged person who rises above the DS index threshold is no longer considered disadvantaged. With respect to value creation by social interventions, the literature on social entrepreneurship has produced evidence that social interventions primarily raise the mean DS index of *disadvantaged* treatment groups (e.g., Austin et al., 2006; Colby et al., 2004; Martin & Osberg, 2007; Nicholls, 2008; Phills et al., 2008; Reed et al., 2005; see also the section entitled "Not-for-Profit Organizations and Social Interventions" in this article).

In conclusion, a social intervention's effectiveness equals the changes that it has initiated in its treatment group's mean DS index, which is likely to lie below the DS index threshold. Thus, an intervention's social character is determined by its treatment of individuals with a DS index below the DS index threshold. For example, the LS threshold of Bangladesh was 5.8 in 2002 (www.worldvaluessurvey.org/WVSONline.jsp, wave 4: v81). An intervention operating in Bangladesh would be considered a social one only if it treated people with an LS below 5.8. Because the social element pertains exclusively to activity below the DS index threshold, a treatment group's social need equals, mathematically speaking, the difference between the DS index threshold and the treatment group's DS index (see Figure 2).

Requirement 3: Control for Different Socioeconomic and Institutional Contexts

Our third requirement for the comparison of the social value creation of two independent interventions is an integration of socioeconomic and institutional aspects. Different living

FIGURE 2
Two Levels of the DS Index Before a Social Intervention Has Taken Place ($t = 0$)



Note: Social need equals the difference between the DS index threshold and the mean DS index of the treatment group.

standards in different regions or countries cause the same absolute improvement in a treatment group's DS index to result in growth rates that are higher in poor countries than in rich ones. In the functional model effectiveness is the degree to which an organization's output benefits society. In terms of degree, then, effectiveness is a relative, not an absolute, construct (Lewin & Minton, 1986). This understanding suggests the need to express social value creation, too, in relative terms. With effectiveness understood as a relative construct, social value creation is reflected in the degree to which an intervention reduces its treatment group's relative social need. Accordingly, this degree equals the improvement of the absolute mean DS index of an intervention's disadvantaged treatment group divided by the social need of that group (see Figure 3). In summary, it should be possible to evaluate an intervention's social value creation in proportion to social need.

Designing a Framework for Comparing Social Value Creation

Progressing from the considerations discussed above, we arrive at the following six variables from the functional model: (1) the mean DS index of the treatment group at $t = 0$, (2) the regional or national DS index threshold, (3) social need in a region or country at $t = 0$, (4) the treatment group's improved mean DS index at $t = 1$, (5) the degree of social value created, and (6) the weighted degree of social value created. We illustrate each variable by calculating the

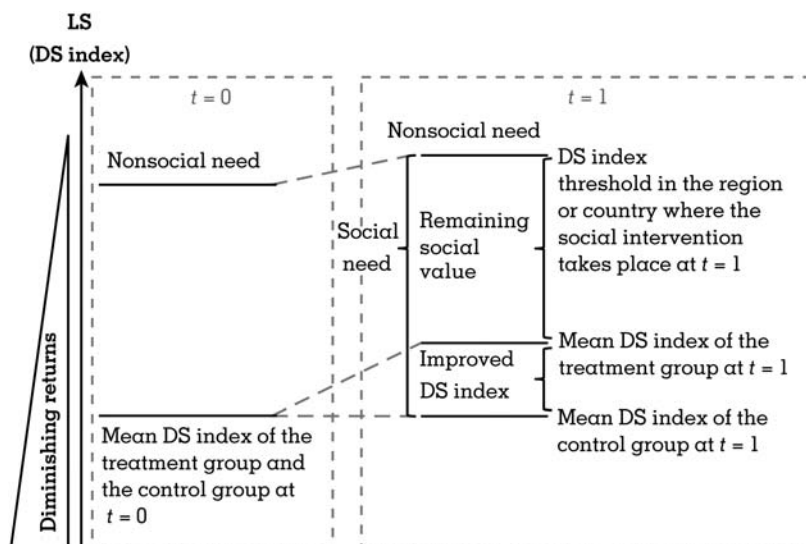
corresponding numbers for an example of a real mental health intervention in the United States described by Suldo, Savage, and Mercer (2014).²

Mean DS index of the treatment group at $t = 0$. The expression $t = 0$ marks the starting point of the social intervention. A treatment group consists of individuals who are directly treated by a social intervention. As outlined in the section "Requirement 1: Uniform Social Value Construct and Measurement Unit," one calculates a treatment group's mean DS index by adding up the DS indices of a representative sample of treatment group members and dividing that sum by the number of treatment group members in the sample at $t = 0$. Comparison of social value creation is possible only if the DS index is standardized for the particular treatment groups that are the targets of the compared interventions. With respect to our example, Suldo et al. (2014) asked a group of twenty-eight disadvantaged students to rate their LS on a scale from 1 (*strongly disagree*) to 6 (*strongly agree*). To evaluate the students' mean LS against the national LS of the United States later, we convert the 6-point Likert-Scale into a 10-point scale. On average, the students reported an LS of 6.7 at the beginning of the intervention.

Regional or national DS index threshold. A regional or national DS index threshold is calculated as the sum of the DS indices of a representative sample of individuals of the focal region or country divided by the number of individuals in the sample. Depending on the country in which a social intervention has taken place, regional or national data are already received from national and supranational organizations, such as the OECD (data on the perceived LS in a country) or the Gallup Healthways Well-Being Index (data related to life evaluation in different U.S. states). For the purpose of our example, we note that the World

² We use the study by Suldo et al. (2014) to illustrate our framework. For the sake of simplicity, however, we assume that the control group has the same initial LS as the treatment group. We note also that the small size of the sample in Suldo et al. (2014) is not statistically representative. Nevertheless, this statistical limitation can be disregarded since the example's purpose is to help us demonstrate the applicability of our framework by social interventions. Studies that already apply LS ratings in organizational settings are very rare.

FIGURE 3
Three Levels of the DS Index



Note: Degree of social value created for one period equals the improved DS index divided by the social need.

Value Surveys (www.worldvaluessurvey.org/WVSONline.jsp, wave 5: V22), Abdallah, Thompson, and Marks (2007), and Veenhoven (2009) report, on average, a mean LS of 7.4 for the United States.

As defined above, a DS is the result of an individual's perceived discrepancy between his or her aspired and achieved levels of need in one area of life. It might be helpful to clearly distinguish an individual's aspired level of need from a regional or national DS index threshold. First, the threshold is an aggregation of the DS indices, whereas the aspired level of need is one determinant for the DS index. Second, an individual's aspired level of need might lie above a regional or national DS index threshold. As a result, a regional or national DS index threshold might influence an aspired level, but a threshold is not the same thing as an aspired level.

Because the effectiveness of a social intervention is assessed from the perspective of the treatment group, the DS index threshold refers to the region or country in which a social intervention takes place, not the region or country in which it originates. If a social enterprise is active in several regions or countries, impact analysts must calculate the DS index threshold for each region or country separately and conduct an impact analysis by region or

country. In keeping with the conclusions of the preceding section, the regional or national DS index threshold reflects the transition between advantaged and disadvantaged individuals. Because we found sound evidence that social interventions usually target disadvantaged individuals, the DS index threshold also determines the threshold between social and nonsocial interventions.

Social need in a region or a country at $t = 0$. The analysis of the framework centers on treatment groups below the regional or national DS index threshold. At $t = 0$ a treatment group's social need can be described as the difference between the previous two variables (the regional or national DS index threshold and the mean DS index of the treatment group). These three variables are illustrated in Figure 2. In our example the social need of the students treated by the mental health intervention was $7.4 - 6.7 = 0.7$.

Treatment group's improved mean DS index at $t = 1$. Assessment of this variable takes place at $t = 1$. The expression $t = 1$ describes a point in time at which the social value created by an intervention is measured, but it does not necessarily refer to the end of the intervention. A treatment group's improved mean DS index might be initiated by the focal social intervention or by effects outside its area of

influence, such as improvements caused by other social interventions, governmental actions, and socioeconomic trends (Mayne, 2001; White, 2010). To control for these external effects, program evaluators adopt experimental and quasi-experimental research designs, such as RCTs or the difference-in-differences technique (e.g., Campbell, 1969; Duflo & Kremer, 2008; Epstein & Klerman, 2012; Meyer, 1995; White, 2013). These research designs are guided by the following counterfactual principle: “think of a cause of something that makes a difference, and the difference it makes must be a difference from what would have happened without it” (Lewis, 1972: 557). Program evaluators calculate this difference by first creating a control group of individuals who have characteristics similar to those of the treatment group but did not receive the social intervention (Sorensen & Grove, 1977) and then comparing this control group’s living conditions with those of the treatment group at $t = 1$ (White, 2013). As with the treatment group’s mean DS index, one calculates the control group’s mean DS index by adding up the DS indices of a representative sample of *nontreated* individuals and then dividing the resulting sum by the number of nontreated individuals in the sample. The difference between the mean DS indices of the treatment and the control group indicates that the social intervention has potentially caused some effect on the treatment group at $t = 1$. Hereafter we term this difference *treatment group’s improved DS index*. In the study by Suldo et al. (2014), a group of fifty-five disadvantaged students was randomly split into a treatment group of twenty-eight people who received a “10-week group wellness-promotion” (Suldo et al., 2014: 19) and a control group of twenty-seven people who received no mental aid services. At $t = 1$ the mean LS of the treatment group was 7.3, whereas the mean LS of the control group was 6.7. Thus, the treatment group’s improved LS would be $7.3 - 6.7 = 0.6$.

Degree of social value created. In the section titled “Requirement 3: Control for Different Socioeconomic and Institutional Contexts,” we concluded that effectiveness is a relative construct. Social value creation is reflected by the degree to which an organization reduces a treatment group’s relative disadvantage. A treatment group’s social need equals the maximum possible social improvement in its mean DS index. At $t = 1$ a treatment group’s social need equals the

difference between the regional or national DS threshold at $t = 1$ and the mean DS index of the control group at $t = 1$. The degree of social value created at $t = 1$ indicates the degree to which a social intervention has contributed to satisfying a treatment group’s need. If a social intervention improves the mean DS index of the treatment group relative to the mean DS index of the control group at $t = 1$, and if the group’s mean DS index has not yet crossed the applied regional or national threshold, then social value has been created. The embeddedness of the treatment group’s improved DS index within its overall social need is illustrated as shown in Figure 3.

With social value creation defined in relative terms, the degree of social value created in the period from $t = 0$ to $t = 1$ equals a treatment group’s improved DS index divided by the absolute social need at $t = 1$. The degree of social value created³ is calculated as

$$\text{degree of SVC}_{i,t} = \frac{\text{DS index improved}_{i,j,t}}{\text{social need}_{i,j,t}}, \quad (1)$$

with i = intervention, $\varepsilon = \{i = 1, 2, \dots, I\}$; t = point in time, $\varepsilon = \{t = 1\}$; j = treatment group, $\varepsilon = \{j = 1, 2, \dots, J\}$.

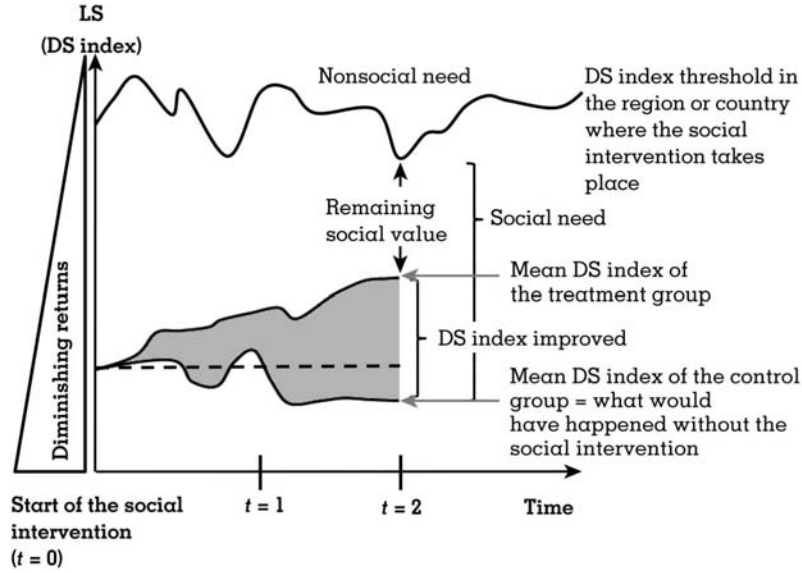
Earlier in this article we discussed possible structures of the DS index. If one aggregates DS by using a constant elasticity of substitution function (Rojas, 2006) or a model containing quadratic and interaction terms (Gonzalez et al., 2010), the marginal returns of the DS index diminish as DS increases. Consequently, improvement in the DS index tends to be easier for high social needs (lower levels of individuals’ DS index) than for low social needs (higher levels of individuals’ DS index) and becomes ever smaller as the relative disadvantage of the treatment group decreases in comparison with the regional or national DS index threshold. These diminishing marginal returns are represented in Figure 3 by a right triangle.

An analysis of the degree of social value created might be relevant for more than one period of time (Figure 4).

If there is *more than one period*, the degree of social value created is calculated by taking the

³ For readability we use the abbreviation “SVC” for “social value created” in the equations.

FIGURE 4
Framework for Comparing Social Value Creation



Note: Degree of social value created for more than one period equals the difference between the degree of social value created at $t + 1$ and the degree of social value created at t .

difference between the treatment group's improved DS index at the beginning and end of a given time span and dividing it by the corresponding social need:

$$\text{degree of } SVC_{i,(t+1-t)} = \text{degree of } SVC_i(t+1) - \text{degree of } SVC_i(t), \quad (2)$$

$$\text{degree of } SVC_{i,(t+1-t)} = \left(\frac{\text{DS index improved}_{ij,t+1}}{\text{social need}_{ij,t+1}} \right) - \left(\frac{\text{DS index improved}_{ij,t}}{\text{social need}_{ij,t}} \right), \quad (3)$$

with i = intervention, $\varepsilon = \{i = 1, 2, \dots, I\}$; t = point in time, $\varepsilon = \{t = 1, 2, 3, \dots, T\}$; j = treatment group, $\varepsilon = \{j = 1, 2, \dots, J\}$.

In our example the satisfaction threshold of the United States is still 7.4. Thus, the degree of social value created by the mental health intervention is calculated as

$$\text{degree of } SVC_{i,(t+1-t)} = \left(\frac{(7.3-6.7)}{(7.4-6.7)} \right) - \left(\frac{(6.7-6.7)}{(7.4-6.7)} \right) = 86 \text{ percent}. \quad (4)$$

Figure 4 also illustrates the interplay between the mean DS indices of the treatment and control groups. An increase in the control group's mean DS index indicates that the treatment group also benefits from external effects (e.g.,

other social interventions or socioeconomic effects). This increase in the control group's mean DS index thus reduces the treatment group's improved DS index. The increase in the control group's mean DS index, however, could exceed any increase in the treatment group's mean DS index. This possibility arises if the treatment group's mean DS index diminishes more rapidly than that of the control group or if the treatment group's mean DS index diminishes while the control group's mean DS index improves. In either case social intervention adversely affects the mean DS index of the treatment group (Bjurulf, Vedung, & Larsson, 2013; Lingane & Olsen, 2004).

An example of negative social value creation would be an intervention granting microcredits to its treatment group without educating the group on how to invest them. The members of the treatment group might simply consume the received money. They would be unable to reimburse the microcredits and would incur debts. In that case the focal treatment group's social need at $t = 1$ would no longer be indicated by the difference between the regional or national DS index threshold and the mean DS index of the control group but, rather, by the difference between the regional or national DS index threshold and the

mean DS index of the *treatment group*. Accordingly, the degree of social value created would be calculated as

$$\text{degree of SVC}_{i,(t+1-t)} = \left(\frac{(\text{DS index improved}_{i,j,t+1})}{(\text{remaining social impact}_{i,j,t+1})} \right) - \left(\frac{(\text{DS index improved}_{i,j,t})}{(\text{remaining social impact}_{i,j,t})} \right), \quad (5)$$

with i = intervention, $\varepsilon = \{i = 1, 2, \dots, I\}$; t = point in time, $\varepsilon = \{t = 1, 2, 3, \dots, T\}$; j = treatment group, $\varepsilon = \{j = 1, 2, \dots, J\}$.

Having elaborated on how to calculate the degree of social value created by one social intervention, we continue by comparing that degree with the degree of social value created by another social intervention.

Weighted degree of social value created. It is highly probable that the treatment groups who benefit from the compared interventions differ in size, so the degree of social value created should be weighted by the number of individuals who benefit from a particular intervention:

$$\begin{aligned} \text{weighted degree of SVC}_{i,(t+1-t)} &= w_{i,j,t+1} * \left(\frac{(\text{DS index improved}_{i,j,t+1})}{\text{social need}_{i,j,t+1}} \right) \\ &- w_{i,j,t} * \left(\frac{(\text{DS index improved}_{i,j,t})}{\text{social need}_{i,j,t}} \right), \quad (6) \end{aligned}$$

with i = intervention, $\varepsilon = \{i = 1, 2, \dots, I\}$; t = point in time, $\varepsilon = \{t = 1, 2, 3, \dots, T\}$; j = treatment group, $\varepsilon = \{j = 1, 2, \dots, J\}$; w = weighting = number of beneficiaries, $\varepsilon = \{w = 1, 2, \dots, W\}$.

With our example of the mental health intervention, the weighted degree of social value created would be calculated as follows:

$$\begin{aligned} \text{weighted degree of SVC}_{i,(t+1-t)} &= 28 * \left(\frac{(7.3-6.7)}{(7.4-6.7)} \right) \\ &- 28 * \left(\frac{(6.7-6.7)}{(7.4-6.7)} \right) = 28 * 0.86 = 24.1. \quad (7) \end{aligned}$$

To compare the weighted degrees of social value created by different social interventions, one must stipulate a time span within which to make the comparison. For this time span the weighted degrees of social value creation must be calculated separately for each social intervention. The separate calculations are necessary because the two interventions might have involved different socioeconomic contexts and

treatment groups. Because evaluation of the social value created would be referenced to the regional or national DS index threshold and the mean DS index of the control group, the framework explicitly does not stipulate the absolute amount by which the treatment group's mean DS index has improved within the defined time span. Instead, the framework suggests that the improved DS index of the treatment group should be contextualized within the social need of a particular disadvantaged treatment group in a particular geographic context (see Figure 4). In our example the mental health intervention might be compared with a fictitious microfinance intervention in Nicaragua. For this example we consult a study by Cox (2012), who assessed the mean LS of 5.1 of dump dwellers in Nicaragua at $t = 1$.⁴ We add the fictitious LS of a control group at $t = 1$, which might be 3.3. The satisfaction threshold of Nicaragua at $t = 1$, according to Abdallah et al. (2007), is 6.5. The number of dump dwellers interviewed by Cox (2012) was thirty. As a result, the weighted social value created of the intervention in Nicaragua is

$$\begin{aligned} \text{weighted degree of SVC}_{i,(t+1-t)} &= 30 * \frac{(5.1-3.3)}{(6.5-3.3)} \\ &= 16.9. \quad (8) \end{aligned}$$

Comparing the mental health intervention in the United States with the microfinance intervention in Nicaragua, the mental health intervention has a higher impact, despite having lower improved LS than the microfinance intervention.

DISCUSSION

In this article we set out to develop a conceptual framework for the comparison of social value. To do so we built on insights from several relevant streams of literature: (1) SWB and life satisfaction, (2) not-for-profit management and

⁴ Cox (2012) calculated a mean LS of 3.6 on a 7-point Likert scale, which we converted to 5.1 on a 10-point Likert scale. However, the small sample sizes limit comparison between this case and the mental health example. Additionally, Suldo et al. (2014) measured LS with six items and a 6-point scale ranging from 1 (*strongly disagree*) to 6 (*strongly agree*), whereas Cox (2012) did so with five items and a 7-point scale ranging from 1 (*extremely dissatisfied*) to 7 (*extremely satisfied*). With regard to the previous footnote, this statistical limitation, too, can be disregarded for our purposes.

social entrepreneurship, (3) program evaluation, and (4) organizational effectiveness. The various components of our framework contain elements of each of these literature streams and combine them in a new way, creating a way to compare different social interventions serving different needs of different treatment groups in different socioeconomic and institutional contexts. It thereby answers various calls to respond to the requirements of such frameworks and clearly fills a hitherto persistent gap in the not-for-profit management and social entrepreneurship literature. Offering appropriate methodologies to quantify and compare social value creation, our framework overcomes the drawbacks of current models and contributes to different streams of literature in several ways.

Uniform Social Value Construct (to Compare Different Treatment Groups and Needs)

First, our framework contributes to the literature on not-for-profit management and social entrepreneurship by showing the utility of applying LS ratings as a statistically proven performance measure in organizational settings. In the economic context, LS is used primarily to evaluate the living standard of different countries to "advance the policymaking process," among other things (Kroll & Delhey, 2013: 13; see also Diener et al., 2013; Petrucci & D'Andrea, 2002; Rossouw & Pacheco, 2012). However, there has been little application of LS to assess the performance of an organization and to compare it with the performance of other organizations (e.g., Becchetti, Castriota, & Solferino, 2011). We show this avenue to be a promising one, using our framework to introduce LS as a uniform social value construct that aligns the measurements of social interventions and subsumes all the different needs of the various treatment groups to which social interventions minister. We suggest considering LS ratings as a statistically proven performance measure (Diener et al., 2013; Kroll & Delhey, 2013) that organizations can apply across different treatment groups and different sectors. We also suggest applying DS indices as the approximation of LS because a DS index strengthens the link between a treatment group's particular needs and its LS ratings. This approach increases accuracy when it comes to attributing a social intervention's outputs to the heightened LS of the intended treatment group (Cummins, 1996).

The previously mentioned high validity and reliability of subjective constructs might be accompanied by downsides. One of them could be that analysts collect satisfaction data by using different scales and methods, such as "Cantril's ladder-of-life" (Kroll & Delhey, 2013: 17) and the "Delighted-Terrible Scale" (Diener, 1984: 546; see also Andrews & Withey, 1974), a practice that might cause statistical inaccuracy. Ideally, comparisons of social interventions build on identical methods of data collection. There is also a risk that "individuals can adapt to misfortune and bad conditions" (Binder, *in press*: 6; see also Diener & Lucas, 2000, and Diener & Suh, 1997). When disadvantaged individuals do so in a life situation that is persistently miserable, they tend to report LS that is better than it actually is. For instance, the longer paraplegics have had their disability, the higher the LS they might report.

This adaption to misfortune might not characterize all life situations. Diener and Lucas (2000), for instance, documented cases in which individuals did not adapt to miserable life situations. According to Schyns, adaption processes that may exist do not impede comparison of DS indices among countries, "since adaption processes are more or less the same everywhere" (1998: 19). Furthermore, in a DS index respondents reflect on each domain separately. An observed adaption effect might not occur for all domains equally, because not all DSs may be affected when disadvantaged individuals adapt somewhat to miserable life situations. In those cases a potential adaption effect would not be reflected in its full intensity at the aggregated level of the DS index. The problem of subjective adjustments might also not be highly relevant for people who do not persistently face misfortune and poor living conditions. According to Layard (2010), Oswald and Wu (2010), and Lawless and Lucas (2011), perceptual ratings have been demonstrated to correlate highly with "non-self-report measures" (Sandvik, Diener, & Seidlitz, 1993: 317). Nevertheless, one way to control for the subjective adjustments by such individuals would be to employ complementary assessments of non-self-reported indicators (Diener & Suh, 1997; Rossouw & Pacheco, 2012), which would then increase the evaluation costs of social interventions (Diener et al., 2013; Petrucci & D'Andrea, 2002). Selected non-self-reported

indicators could also control for manipulation of DS index ratings, as when policy makers or social interveners themselves might be tempted to influence respondents to report higher satisfaction than what those respondents actually perceive (Binder, in press).

Social Element (to Determine the "Social" Character of Social Value Creation)

Second, an important contribution that this research makes to the not-for-profit management and social entrepreneurship literature is that it provides a new conceptualization of the "social" element that is at the core of every discussion of social value creation. Given this key element, our research supplements various discussions, such as helping to specify the role of social interventions in society. We acknowledge that social interventions primarily aim to serve people below the LS threshold of a region or a country. The treatment of people below this threshold thus constitutes the "social" character of social interventions. Past not-for-profit management literature reveals a lack of a common understanding and definition of the social element (Nicholls, 2008; Nicholls & Cho, 2008; Polonsky & Grau, 2011). Our framework helps fill this gap.

By defining the social character of social interventions, our framework also contributes clarity to the discussion on an appropriate definition of the "essentially contested" (Choi & Majumdar, 2014: 363) social entrepreneurship construct (e.g., Dacin et al., 2010; Weerawardena & Mort, 2006; Zahra et al., 2009). We conclude that social entrepreneurship focuses on activities primarily benefiting individuals below the threshold of the respective region or country. Changes in a treatment group's mean DS index might very well lead to changes in the DS index threshold, depending on the size of the treatment group. Critical scholars might argue that it is theoretically possible to pretend to create social value by simply reducing the DS index of advantaged individuals. However, our definition of social value creation precludes such reduction, because social value is created only if it is the DS index of disadvantaged people that is primarily raised. "Primarily" means that the absolute value creation for disadvantaged people exceeds the absolute

value creation or reduction for advantaged people.

Comparing Social Interventions Operating in Different Socioeconomic and Institutional Conditions

Third, we contribute to the not-for-profit management and social entrepreneurship literature noting the importance of considering, for social interventions, the socioeconomic and institutional conditions in a region or country (e.g., Austin et al., 2006; Dorado & Ventresca, 2013; Kerlin, 2006; Mair & Marti, 2009; Mair, Marti, & Ventresca, 2012; Miller et al., 2012). Although affording valuable insights into institutions that affect the operations and emergence of social interventions, these studies concentrate mostly on the different contextual effects themselves. They leave entirely unsolved the issue of how to control for socioeconomic and institutional conditions when comparing social interventions in different regions or countries. Other authors conceptualizing social value creation have ignored contextual effects altogether (e.g., Nicholls & Cho, 2008; Swanson & Di Zhang, 2010; Weerawardena & Mort, 2006). With our framework we are able to fill this tremendous research gap. Our suggestion is to evaluate the improved DS index against the background of the social need determined by the socioeconomic and institutional conditions in the given region or country. The social need, in turn, is defined as the extent of a treatment group's *relative* disadvantage within a region or country, which is expressed by the absolute difference between the DS index threshold of this region or country and the treatment group's mean DS index. Our framework contributes a theoretically grounded, simple, and meaningful approach to comparing different social interventions from different socioeconomic and institutional contexts.

By contextualizing the improved DS index within the social need, we also add to the literature on program evaluation. Most researchers have only either referred to value creation as the treatment group's *absolute* improved living conditions (e.g., Duflo & Kremer, 2008; Meyer, 1995) or indicated the progress rate of the treatment group's living conditions (e.g., Banerjee, Duflo, Glennerster, & Kinnan, 2010; Crepon, Devoto, Duflo, & Pariente, 2011). However, comparing progress rates might exaggerate the

effectiveness of a social intervention in low-number ranges and thereby result in misleading conclusions (see our third requirement). Our framework overcomes this limitation. We suggest analyzing social interventions' value creation not in absolute terms only (by only assessing the treatment group's improved mean DS index at $t = 1$) but also relative to the treatment group's need in the corresponding region or country where the respective social intervention operates. With our framework we therefore do not ask how great a percentage of the well-being of the students improved but, rather, what degree of their social need was satisfied by the social intervention. In our example of the mental health intervention (Suldo et al., 2014), an improved LS of 0.6 within a social need of 0.7 indicates that the intervention was able to solve the need by 86%, whereas in our second example a treatment group's improved LS of 1.8 within a fictitious social need of 3.2 would satisfy only 56% of the need.

However, by including experimental and quasi-experimental analyses in our framework to calculate the treatment group's improved mean DS index at $t = 1$, we find that our framework is also subject to the drawbacks of such research designs. RCTs require a target population that is generally much larger than the treatment group of a social intervention. Epstein and Klerman (2012) recommend that a target population be at least twice as large as the studied treatment group. To achieve statistical power, the treatment group, too, must be sufficiently large. This requirement limits the applicability our approach has for small- and maybe even medium-size interventions, which simply might not treat enough people to achieve the necessary sample size. Some critics have also noted that social interventions often lack the time and other resources to conduct RCTs (e.g., Figueredo & Sechrest, 2001; Holma & Kontinen, 2011). Holma and Kontinen have ascertained that, for many social interventions, "experimental evaluations are usually beyond the limit of their capabilities, due to a lack of resources and competence" (2011: 184). This argument has been qualified by Duflo and Kremer (2008). Having carried out numerous RCTs, these authors maintain that such studies are "labor intensive and costly, but no more so than other data collection activities" (2008: 107). In addition, our framework builds on self-reported perceptual

measures, which are "relatively inexpensive and easy to administer" (Diener et al., 2013: 498). According to Tay et al. (2013), both the complexity and length of surveys—even ones consisting of multi-item measures, such as the DS index—are significantly lower than those of other evaluation methods. Bamberger, Rugh, Church, and Fort have asserted that "considerable savings in costs and time can often be achieved through reducing the length and complexity of the survey instrument" (2004: 18).

Another way for social interventions to reduce the sample sizes and costs of evaluation is to compromise statistical power (e.g., Lipsey, 1990). This alternative could be the most practical and, often, the most sufficient approach for interveners unable or unwilling to fulfill the requirements for an RCT study (Bamberger et al., 2004). For instance, if an RCT study is unsuitable for evaluating treatment groups' improved DS index, then quasi-experimental designs such as propensity score matching or the difference-in-differences technique could offer alternative solutions (Duflo & Kremer, 2008; Meyer, 1995; White, 2013). Quasi-experimental designs use "statistical means to construct a comparison group, which, like the control group in an RCT, has the same characteristics as the treatment group" (White, 2013: 35). As Duflo and Kremer (2008) have noted, concerns surrounding quasi-experimental designs include the omission of variables and the white-washing of results. They have concluded that "nonrandomized evaluation methods must be tackled with extreme care" (2008: 98). A before-and-after research design, which makes a control group dispensable, might also be an option if one "has strong evidence that the two groups would have been comparable over time in the absence of the treatment" (Meyer, 1995: 154). White states "that attribution statements, causality statements, in many areas of science are made without a counterfactual," as in "the analysis of the moon causing tides" (2010: 158). For social interventions with simple operational models, a before-and-after research design might approximate treatment groups' improved DS index.

In conclusion, measuring a treatment group's improved DS index might prove to be challenging and could require caution in evaluation. Our framework, too, is subject to this constraint. However, there are several well-proven research

designs that allow sufficient measurement of the treatment group's improved DS index under the conditions that we have described. We therefore feel confident that, even at this juncture, our framework is able to mirror the respective changes in treatment groups' DS indices. Duflo and Kremer (2008: 93) have argued that it is possible to expand the use of randomized evaluations liberally. Further development of appropriate research designs is, thus, both probable and desirable. Institutional funders and politicians are under increasing pressure to foster research on evaluation methods, as shown by *Evaluation's* 2012 special issue on contribution analysis; the existence of sixteen European master's degree programs exclusively dedicated to evaluation courses; the "growing use, notably amongst economists and political scientists, of a range of [evaluation] approaches" (White, 2013: 30); and the growing number of evaluation studies financed by development banks (White, 2010).

Autonomy of the Functional Model (to Base the Framework On)

Fourth, we contribute to the organizational effectiveness literature by demonstrating the appropriateness, independence, and legitimacy of the functional model for the specific social context. With respect to our framework for comparing social value creation, we conclude that the effectiveness of a social intervention is determined by the degree to which it is able to serve the needs of its treatment group. Unlike other effectiveness models, such as the goal or the resource model, the usefulness of the functional model lies in its ability to capture the effectiveness of social interventions: the functional model determines the goal of social interventions to benefit society. It notes the importance that socioeconomic and institutional conditions have for social interventions (Parsons, 1956: 66; Webb, 1974: 665; see also Connolly et al., 1980). It defines effectiveness as a relative, not an absolute, construct (Lewin & Minton, 1986), allowing us to express the amount of social value created equally in relative terms. By demonstrating the utility of the functional model for our framework, we expand our current knowledge on organizational effectiveness theory in general and its

applicability to the not-for-profit context in particular.

Implications for Theory and Research

Our framework has several key theoretical implications. First and most important, our work advances the literature on not-for-profit management and social entrepreneurship by presenting an approach that allows us to compare social value creation of entirely unrelated social interventions operating in different regions or countries and serving different treatment groups. It might have the potential to challenge the current dominant paradigm, according to which it is impossible to compare social value creation of heterogeneous interventions. Disassociated from the logic of commercial value creation, our approach differs significantly from what is currently available in this social field of research (e.g., Bagnoli & Megali, 2011; Diochon & Anderson, 2009; McLoughlin et al., 2009; Swanson & Di Zhang, 2010), and it adds an entirely new perspective to the dominant rationale. Our work thus suggests that disassociating oneself from the logic of commercial value creation and applying our framework, or any other new perspective, is a promising approach for future research in this field of measurement and comparison in the not-for-profit sector.

Moreover, our work has an implication for the ongoing debate on the definition of what "social" means. Our research advances the understanding of what is social in the social sector. This understanding might help sharpen typologies intended to distinguish social entrepreneurship from related types of entrepreneurship (e.g., Dacin et al., 2010; Harris, Sapienza, & Bowie, 2009; McMullen, 2010; Short, Moss, & Lumpkin, 2009; Swanson & Di Zhang, 2010). A more precise understanding of social entrepreneurship "will extend our understanding of this valuable phenomenon and facilitate the development of managerial strategies to assist those who undertake social enterprises" (Dacin et al., 2010: 53).

Our new understanding of social might also be applicable to defining the meaning of the "essentially contested concept" (Gjoeldberg, 2009: 12) of corporate social responsibility (CSR). Clarkson asserted that the "failure, together with the confusion and misunderstanding about the definition and meaning of corporate social

responsibility, ... is a direct result of the inclusive and vague meaning of the word social" (1995: 102). According to our definition of "social," our findings suggest that socially responsible corporations predominantly benefit their disadvantaged stakeholders or disadvantaged individuals in society. This finding might rapidly change the discourse on CSR in the relevant literature. Future research could investigate the extent to which our understanding of social is adaptable to the concept of CSR.

Our framework also has implications, on the "lack of feasible measurement methods" (Costa & Menichini, 2013: 152), for measuring CSR activities among different countries. Scholars assert that these national differences in CSR activities are due to "long-standing, historically entrenched institutions" (Matten & Moon, 2008: 406; see also Gjoeldberg, 2009; Kim, Amaeshi, Harris, & Suh, 2013). Hence, measurement methods "must be able to measure CSR practices across different political-economic systems" (Gjoeldberg, 2009: 11). Our framework suggests that the national satisfaction threshold, which is a result of the socioeconomic and institutional conditions in a region or country, might make it possible to embed the respective CSR activities in their particular context. The perceptual character of satisfaction ratings—national or otherwise—is also in line with Gallardo-Vázquez and Sanchez-Hernandez's work; these authors call perception ratings an "original, valid, and reliable measure of CSR" (2014: 5). Given this statement, our framework also has implications for the search for appropriate indicators with which to measure CSR activities. Turker (2008) and Costa and Menichini (2013) have demonstrated the applicability of stakeholders' perceptual ratings of company commitment. Our framework goes one step beyond because stakeholders, the treatment groups of social interventions, rate not only an organization's social commitment but also the impact of the organization's social activities on their own well-being. Future research should investigate the adaptability of LS ratings as an indicator for corporate social performance.

Our framework for comparing social value creation has implications for the literature on SWB, too. Very few quantitative studies exist that use LS ratings for evaluating the impact of a social intervention (e.g., Becchetti et al., 2011; Suldo et al., 2014). The framework demonstrates

that these studies point in a promising direction and calls for more empirical studies of this kind.

An important avenue for future research lies in empirically testing this framework. We hope that future scholars can elaborate on our guidance and suggestions for operationalizing our framework and can conduct initial empirical studies that build on it. In this context we support Duflo and Kremer's (2008) call for research on evaluation designs. In addition to comparative studies on experimental and quasi-experimental design, progress on evaluation techniques that can resolve issues with attribution and contribution would widen the applicability of our framework. We argue that program evaluations should assess a social intervention's value creation relative to the social need of the particular treatment group in the given region or country. This relational assessment is especially relevant for evaluating the external validity of replicated programs in other contexts (Stame, 2010; Tilley, 1996). The authors of the current literature on program evaluation assume that the impact of social interventions can be generalized if evaluation shows similar absolute results in different settings (e.g., Bold, Kimenyi, Mwabu, & Sandefur, 2013; Duflo, Glennerster, & Kremer, 2007; Woolcook, 2013). However, it might be valuable to test the currency of this assumption by relating the results of respective studies to the social need in the country. Contextualizing the absolute impact of a social intervention within the social need allows one to enter into context-specific discussions about the relative merits of (and priority that should be accorded to) roads, irrigation, cash transfers, immunization, legal reform etc with some degree of grounded confidence—i.e., on the basis of appropriate metrics, theory, experience and ... trajectories of change (Woolcook, 2013: 231).

Our work has implications for the theory of organizational effectiveness. We contribute to this stream of literature by demonstrating both the autonomy of the functional model in the not-for-profit context and its clear legitimacy in the field of organizational effectiveness research. Because social enterprises have special characteristics and a specific context, they primarily pursue the organizational purpose postulated in the functional model: to benefit society. Unlike the case in the commercial context, ensuring one's own viability is seen as a means to an end in the not-for-

profit context. In the particular not-for-profit context, the functional model is even more appropriate than all the other well-established multivariate organizational effectiveness models. We therefore suggest reintroducing the functional model as an independent and autonomous alternative in the organizational effectiveness literature and stress that the relevance of the functional model be strengthened in the literature on both organizational effectiveness in general and not-for-profit organizations in particular. This view is shared by Cameron and Whetten, who stated that "multiple viewpoints all may be equally legitimate but under different circumstances and with different types of organizations" (1983: 274). Having demonstrated that the functional model—as a hybrid of established organizational effectiveness models—can be considered an autonomous approach in the social context, we assume that there is room for more independent hybrid forms that are particularly suitable for specific "niches." It is a view supported by Cameron and Whetten's early suggestions to "recognize the composited nature of the construct" (1983: 275) of organizational effectiveness and to systematically integrate existing models with one another. Extending their argument, we suggest that further research explore additional hybrid types of effectiveness models that might be able to integrate certain currently unconnected characteristics of original models and serve certain sectors or specific contexts more appropriately than do traditional original models.

Implications for Practice

Our framework for comparing social value creation may affect various stakeholder groups of social interventions, particularly managers of not-for-profit organizations and social enterprises, impact investors, corporations, and governments. For managers of not-for-profit organizations and social enterprises, cross-sector and cross-national comparability might increase competition between social interveners (McLoughlin et al., 2009). Social enterprises and other social organizations already "face increasing pressures to demonstrate the results of their work both from donors and from [the] public" (Holma & Kontinen, 2011: 181). These expectations could even grow if our framework is operationalized and applied. Although some

social enterprises and charity organizations might perceive them as unacceptable, others might welcome such demands because managers of not-for-profit interventions often want to benchmark their own intervention's weighted degree of social value creation with other, well-established social interventions (Lingane & Olsen, 2004). Benchmarking with the regional or national DS index threshold also helps distinguish between social and nonsocial activities and thereby partially legitimates an organization's claim that it is primarily a social organization (Bitektine, 2011).

Our framework is also valuable for impact investors, a new type of investor (Nicholls, 2009). Unlike conventional investors, impact investors actively look for the social performance of potential investment alternatives and provide both financial and nonfinancial support to social enterprises (Lingane & Olsen, 2004; Mair & Hehenberger, in press). Because social criteria are crucial in the screening and due diligence processes that impact investors carry out, these actors are particularly interested in comparing the social effectiveness of potential impact investment options (Buckmaster, 1999; Miller et al., 2012). Our framework might aid the investment decisions of impact investors because it facilitates comparison between heterogeneous interventions. It thereby creates a sound basis for one of the biggest challenges in launching a social stock exchange: the issue of choosing "among a bewildering multitude of outstretched hands" (*The Daily Beast*, 2008). Application of cross-sector and cross-national comparisons would enable social stock exchanges to give stockholders a simple and useful decision criterion for investing their money.

Our framework also has implications for practices of corporate social reporting. Commercial enterprises are often accused of embellishing their social activities (Bouten, Everaert, Liedekerke, & De Moor, 2011; Font, Walmsley, Cogotti, McCombes, & Hdusler, 2012). According to Bouten et al., businesses would improve the credibility of their social reporting if they would "elaborate on the fulfillment of ... [their CSR] commitments and the achieved outcomes" (2011: 189). Our framework enables corporations to assess outcomes that benefit disadvantaged stakeholder groups, so it might help enhance the credibility of social reporting. Furthermore, current CSR reporting often pays more attention to

environmental performance than social performance, a focus that may stem from difficulties in assessing social performance (Hahn & Kiihnen, 2013; see also Font et al., 2012). Our framework contributes to solving these difficulties. The focus on social performance might thus increase if companies use our approach. Because our framework also takes into account regional or national socioeconomic and institutional conditions, it might be particularly valuable for multinational corporations. After all, current "reporting practices may ... vary across countries and regions due to different cultural and social norms or governmental regulations" (Hahn & Kiihnen, 2013: 13; see also Golob & Bartlett, 2007).

Last, governments can refine their approach to social problems by subsidizing the most effective social interventions first. The goal of achieving equality of satisfaction across individuals within one region or country might serve as a complement or alternative to the "standard welfare measure of GDP per capita" (Ura et al., 2012: 5; see also Sen, 1985).

CONCLUSION

As noted in the literature review, comparing unrelated interventions has been regarded as impossible. The framework we have presented, however, demonstrates that cross-sector and cross-national comparisons of social interventions are indeed feasible. We thereby contribute to solving the age-old problems associated with comparing unrelated social interventions that serve different needs of different treatment groups in different socioeconomic and institutional contexts. We hope our work encourages other scholars to advance the framework presented here or to develop alternative approaches that promote cross-sector and cross-national comparisons of social value creation.

REFERENCES

- Abdallah, S., Thompson, S., & Marks, N. 2007. Estimating worldwide life satisfaction. *Ecological Economics*, 65: 35–47.
- Anand, P., Hunter, G., Carter, I., Dowding, K., Guala, F., & Van Hees, M. 2009. The development of capability indicators. *Journal of Human Development and Capabilities*, 10: 125–152.
- Andrews, F. M., & Withey, S. B. 1974. Developing measures of perceived life quality: Results from several national surveys. *Social Indicators Research*, 1: 1–26.
- Austin, J., Stevenson, H., & Wei-Skillern, J. 2006. Social and commercial entrepreneurship: Same, different, or both? *Entrepreneurship Theory and Practice*, 30: 1–23.
- Bagnoli, L., & Megali, C. 2011. Measuring performance in social enterprises. *Nonprofit and Voluntary Sector Quarterly*, 40: 149–165.
- Bamberger, M., Rugh, J., Church, M., & Fort, L. 2004. Shoestring evaluation: Designing impact evaluations under budget, time and data constraints. *The American Journal of Evaluation*, 25: 5–37.
- Banerjee, A., Duflo, E., Glennerster, R., & Kinnan, C. 2010. *The miracle of microfinance? Evidence from a randomized evaluation*. Working paper No. 18950, National Bureau of Economic Research, Cambridge, MA.
- Becchetti, L., Castriota, S., & Solferino, N. 2011. Development projects and life satisfaction: An impact study on fair trade handicraft producers. *Journal of Happiness Studies*, 12: 115–138.
- Beja, E. L., & Yap, D. B. 2013. Counting happiness from the individual level to the group level. *Social Indicators Research*, 114: 621–637.
- Berger, P. L., & Luckmann, T. 1967. *The social construction of reality*. New York: Anchor Books.
- Binder, M. In press. Subjective well-being capabilities: Bridging the gap between the capability approach and subjective well-being research. *Journal of Happiness Studies*.
- Bitektine, A. 2011. Toward a theory of social judgments of organizations: The case of legitimacy, reputation, and status. *Academy of Management Review*, 36: 151–179.
- Bjurulf, S., Vedung, E., & Larsson, C. G. 2013. A triangulation approach to impact evaluation. *Evaluation*, 19: 56–73.
- Bold, T., Kimenyi, M., Mwabu, A. N., & Sandefur, J. 2013. *Scaling up what works: Experimental evidence on external validity in Kenyan education*. Working paper No. 321, Center for Global Development, Washington, DC.
- Boschee, J. 2008. Social entrepreneurship: The promise and the perils. In A. Nicholls (Ed.), *Social entrepreneurship: New models of sustainable social change*: 356–391. Oxford: Oxford University Press.
- Bouten, L., Everaert, P., Liedekerke, L. V., & De Moor, L. 2011. Corporate social responsibility reporting: A comprehensive picture? *Accounting Forum*, 35: 187–204.
- Bradburn, N. M. 1969. *The structure of psychological well-being*. Chicago: Aldine.
- Buckmaster, N. 1999. Associations between outcome measurement, accountability and learning for non-profit organizations. *International Journal of Public Sector Management*, 12: 186–197.
- Burdge, R. J., & Vanclay, F. 1996. Social impact assessment: A contribution to the state of the art series. *Impact Assessment*, 14: 59–86.

- Busseri, M. A., & Sadava, S. W. 2011. A review of the tripartite structure of subjective well-being: Implications for conceptualization, operationalization, analysis, and synthesis. *Personality and Social Psychology Review*, 15: 290–314.
- Cameron, K. S. 1986. Effectiveness as paradox: Consensus and conflict in conceptions of organizational effectiveness. *Management Science*, 32: 539–553.
- Cameron, K. S., & Whetten, D. A. 1983. Some conclusions about organizational effectiveness. In K. Cameron & D. A. Whetten (Eds.), *Organizational effectiveness. A comparison of multiple models*: 261–277. New York: Academic Press.
- Campbell, A., Converse, P. E., & Rodgers, W. L. 1976. *The quality of American life: Perceptions, evaluations, and satisfactions*. New York: Russell Sage Foundation.
- Campbell, D. T. 1969. Reforms as experiments. *The American Psychologist*, 24: 409–429.
- Campbell, D. T. 2002. Outcomes assessment and the paradox of nonprofit accountability. *Nonprofit Management & Leadership*, 12: 243–259.
- Carp, F. M., & Carp, A. 1982. Test of a model of domain satisfactions and well-being: Equity considerations. *Research on Aging*, 4: 503–522.
- Chell, E. 2007. Social enterprise and entrepreneurship: Towards a convergent theory of the entrepreneurial process. *International Small Business Journal*, 25: 5–26.
- Choi, N., & Majumdar, S. 2014. Social entrepreneurship as an essentially contested concept: Opening a new avenue for systematic future research. *Journal of Business Venturing*, 29: 363–376.
- Clark, C., Rosenzweig, W., Long, D., & Olsen, S. 2004. *Double bottom line project report: Assessing social impact in double bottom line ventures*. Working paper No. 13, University of California, Berkeley.
- Clarkson, M. B. E. 1995. A stakeholder framework for analyzing and evaluating corporate social performance. *Academy of Management Review*, 20: 92–117.
- Colby, S., Stone, N., & Carttar, P. 2004. Zeroing in on impact: In an era of declining resources, nonprofits need to clarify their intended impact. *Stanford Social Innovation Review*, 2: 24–33.
- Connolly, T., Colon, E. J., & Deutsch, S. J. 1980. Organizational effectiveness: A multiple-constituency approach. *Academy of Management Review*, 5: 211–217.
- Costa, R., & Menichini, T. 2013. A multidimensional approach for CSR assessment: The importance of the stakeholder perception. *Expert Systems with Applications*, 40: 150–161.
- Cox, K. 2012. Happiness and unhappiness in the developing world: Life satisfaction among sex workers, dump-dwellers, urban poor, and rural peasants in Nicaragua. *Journal of Happiness Studies*, 13: 103–128.
- Crepon, B., Devoto, F., Duflo, E., & Pariente, W. 2011. *Impact of microcredit in rural areas of Morocco: Evidence from a randomized evaluation*. Working paper, J-PAL Europe, MIT.
- Cummins, R. A. 1996. The domains of life satisfaction: An attempt to order chaos. *Social Indicators Research*, 38: 303–328.
- Cummins, R. A., Eckersley, R., Pallant, J., Van Vugt, J., & Misajon, R. 2003. Developing a national index of subjective wellbeing: The Australian unity wellbeing index. *Social Indicators Research*, 64: 159–190.
- Cunningham, J. B. 1977. Approaches to evaluation of organizational effectiveness. *Academy of Management Review*, 2: 463–474.
- Dacin, P. A., Dacin, M. T., & Matear, M. 2010. Social entrepreneurship: Why we don't need a new theory and how we move forward from here. *The Academy of Management Perspectives*, 24: 37–57.
- The Daily Beast*. 2008. A stock exchange for do-gooders. May 31: <http://www.thedailybeast.com/newsweek/2008/05/31/a-stock-exchange-for-do-gooders.html>.
- Davis, P. S., & Pett, T. L. 2002. Measuring organizational efficiency and effectiveness. *Journal of Management Research*, 2: 87–97.
- Dees, J. G. 1998. Enterprising nonprofits: What do you do when traditional sources of funding fall short? *Harvard Business Review*, 55–67.
- Denison, D. 1990. *Corporate culture and organizational effectiveness*. New York: Wiley.
- Di Domenico, M. L., Haugh, H., & Tracey, P. 2010. Social bricolage: Theorizing social value creation in social enterprises. *Entrepreneurship Theory and Practice*, 34: 681–703.
- Diener, E. 1984. Subjective well-being. *Psychological Bulletin*, 95: 542–575.
- Diener, E., Inglehart, R., & Tay, L. 2013. Theory and validity of life satisfaction scales. *Social Indicators Research*, 112: 497–527.
- Diener, E., & Lucas, R. E. 2000. Explaining differences in societal levels of happiness: Relative standards, need fulfillment, culture, and evaluation theory. *Journal of Happiness Studies*, 1: 41–78.
- Diener, E., & Suh, E. 1997. Measuring quality of life: Economic, social, and subjective indicators. *Social Indicators Research*, 40: 189–216.
- Diochon, M., & Anderson, A. R. 2009. Social enterprise and effectiveness: A process typology. *Social Enterprise Journal*, 5: 7–29.
- Donaldson, T. 2001. The ethical wealth of nations. *Journal of Business Ethics*, 31: 25–36.
- Dorado, S., & Ventresca, M. J. 2013. Crescive entrepreneurship in complex social problems: Institutional conditions for entrepreneurial engagement. *Journal of Business Venturing*, 28: 69–82.
- Duflo, E., Glennerster, E., & Kremer, M. 2007. *Using randomization in development economics research: A toolkit*. Working paper No. 6059, Centre for Economic Policy Research, London.
- Duflo, E., & Kremer, M. 2008. *Use of randomization in the evaluation of development effectiveness*. Boston: MIT.

- Economics. Available at <http://economics.mit.edu/files/2785>.
- Early, L. T. 2005. Editorial. The philanthropic enterprise: Reassessing the means and ends of philanthropy. *Journal of Economic Affairs*, 25: 2–4.
- Emerson, J. 2003. The blended value proposition: Integrating social and financial returns. *California Management Review*, 45: 35–51.
- Epstein, D., & Klerman, J. A. 2012. When is a program ready for rigorous impact evaluation? The role of falsifiable logic model. *Evaluation Review*, 36: 375–401.
- Erdogan, B., Bauer, T. N., Truxillo, D. M., & Mansfield, L. R. 2012. Whistle while you work: A review of the life satisfaction literature. *Journal of Management*, 38: 1038–1083.
- Esteves, A. M., Franks, D., & Vanclay, F. 2012. Social impact assessment: The state of the art. *Impact Assessment and Project Appraisal*, 30: 34–42.
- Etzioni, A. 1960. Two approaches to organizational analysis: A critique and a suggestion. *Administrative Science Quarterly*, 5: 257–278.
- Figueredo, A. J., & Sechrest, L. 2001. Approaches used in conducting health outcomes and effectiveness research. *Evaluation and Program Planning*, 24: 41–59.
- Finsterbusch, K. 1985. State of the art in social impact assessment. *Environment and Behavior*, 17: 193–221.
- Finsterbusch, K. 1995. In praise of SIA—A personal review of the field of social impact assessment: Feasibility, justification, history, methods, issues. *Impact Assessment*, 13: 229–252.
- Font, X., Walmsley, A., Cogotti, S., McCombes, L., & Hausler, N. 2012. Corporate social responsibility: The disclosure-performance gap. *Tourism Management*, 33: 1544–1553.
- Forbes, D. 1998. Measuring the unmeasurable: Empirical studies of nonprofit organization effectiveness. *Nonprofit and Voluntary Sector Quarterly*, 27: 183–202.
- Foundation Center 2012. *Tools and resources for assessing social impact*. TRASI database: <http://trasi.foundation-center.org/browse.php>.
- Gallardo-Vázquez, D., & Sanchez-Hernandez, M. I. 2014. Measuring corporate social responsibility for competitive success at a regional level. *Journal of Cleaner Production*, 72: 14–22.
- Georgopoulos, B. S., & Mann, F. C. 1962. *The community general hospital*. New York: Macmillan.
- Gjoeldberg, M. 2009. Measuring the immeasurable? Constructing an index of CSR practices and CSR performance in 20 countries. *Scandinavian Journal of Management*, 25: 10–22.
- Golob, U., & Bartlett, J. L. 2007. Communicating about corporate social responsibility: A comparative study of CSR reporting in Australia and Slovenia. *Public Relations Review*, 33: 1–9.
- Gonzalez, M., Coenders, G., Saez, M., & Casas, F. 2010. Non-linearity, complexity and limited measurement in the relationship between satisfaction with specific life domains and satisfaction with life as a whole. *Journal of Happiness Studies*, 11: 335–352.
- Grimes, M. 2010. Strategic sensemaking within funding relationships: The effects of performance measurement on organizational identity in the social sector. *Entrepreneurship Theory and Practice*, 34: 763–783.
- Hahn, R., & Kiihnen, M. 2013. Determinants of sustainability reporting: A review of results, trends, theory, and opportunities in an expanding field of research. *Journal of Cleaner Production*, 59: 5–21.
- Hannan, M. T., & Freeman, J. 1977. Obstacles to comparative studies. In P. S. Goodman & J. M. Pennings (Eds.), *New perspectives on organizational effectiveness*: 106–131. San Francisco: Jossey-Bass.
- Harris, J. D., Sapienza, H. J., & Bowie, N. E. 2009. Ethics and entrepreneurship. *Journal of Business Venturing*, 24: 407–418.
- Henri, J.-F. 2004. Performance measurement and organizational effectiveness: Bridging the gap. *Managerial Finance*, 30: 93–123.
- Herman, R. D. 1990. Methodological issues in studying the effectiveness of nongovernmental and nonprofit organizations. *Nonprofit and Voluntary Sector Quarterly*, 19: 293–306.
- Herman, R. D., & Renz, D. O. 2008. Advancing nonprofit organizational effectiveness research and theory: Nine theses. *Nonprofit Management & Leadership*, 18: 399–415.
- Hervieux, C., Gedajlovic, E., & Turcotte, M. F. B. 2010. The legitimization of social entrepreneurship. *Journal of Enterprising Communities: People and Places in the Global Economy*, 4: 37–67.
- Holma, K., & Kontinen, T. 2011. Realistic evaluation as an avenue to learning for development NGOs. *Evaluation*, 17: 181–192.
- Hsieh, C.-M. 2003. Counting importance: The case of life satisfaction and relative domain importance. *Social Indicators Research*, 61: 227–240.
- ICGP (Interorganizational Committee on Guidelines and Principles for Social Impact Assessment). 1994. *Guidelines and principles for social impact assessment*. http://www.nmfs.noaa.gov/sfa/social_impact_guide.htm.
- ICGP (Interorganizational Committee on Guidelines and Principles for Social Impact Assessment). 2003. Principles and guidelines for social impact assessment in the United States. *Impact Assessment and Project Appraisal*, 21: 231–250.
- Kerlin, J. A. 2006. Social enterprise in the United States and Europe: Understanding and learning from the differences. *Voluntas*, 17: 247–263.
- Keyes, C. L. M. 1998. Social well-being. *Social Psychology Quarterly*, 61: 121–140.
- Kim, C. H., Amaeshi, K., Harris, S., & Suh, C.-J. 2013. CSR and the national institutional context: The case of South Korea. *Journal of Business Research*, 66: 2581–2591.
- Kroll, C., & Delhey, J. 2013. A happy nation? Opportunities and challenges of using subjective indicators in policy-making. *Social Indicators Research*, 114: 13–28.

- Lawless, N. M., & Lucas, R. E. 2011. Predictors of regional well-being: A country level analysis. *Social Indicators Research*, 101: 341–357.
- Layard, R. 2010. Measuring subjective well-being. *Science*, 327: 534–535.
- Lecy, J. D., Schmitz, H. P., & Swedlund, H. 2012. Non-governmental and not-for-profit organizational effectiveness: A modern synthesis. *Voluntas: International Journal of Voluntary and Nonprofit Organizations*, 23: 434–457.
- Lewin, A. Y., & Minton, J. W. 1986. Determining organizational effectiveness: Another look, and an agenda for research. *Management Science*, 32: 514–538.
- Lewis, D. 1972. Causation. *The Journal of Philosophy*, 70: 556–567.
- Lingane, A., & Olsen, S. 2004. Guidelines for social return on investment. *California Management Review*, 46: 116–135.
- Lipsey, M. W. 1990. *Design sensitivity: Statistical power for experimental research*. Newbury Park, CA: Sage.
- Locke, E. A. 1969. What is job satisfaction? *Organizational Behavior and Human Performance*, 4: 309–336.
- Lockie, S. 2001. SIA in review: Setting the agenda for impact assessment in the 21st century. *Impact Assessment and Project Appraisal*, 19: 277–287.
- Mair, J., & Hehenberger, L. In press. Front stage and back stage convening: The transition from opposition to mutualistic co-existence in organizational philanthropy. *Academy of Management Journal*.
- Mair, J., & Marti, I. 2006. Social entrepreneurship research: A source of explanation, prediction, and delight. *Journal of World Business*, 41: 36–44.
- Mair, J., & Marti, I. 2009. Entrepreneurship in and around institutional voids: A case study from Bangladesh. *Journal of Business Venturing*, 24: 419–435.
- Mair, J., Marti, I., & Ventresca, M. J. 2012. Building inclusive markets in rural Bangladesh: How intermediaries work institutional voids. *Academy of Management Journal*, 55: 819–850.
- Mallard, A. G. C., Lance, C. E., & Michalos, A. C. 1997. Culture as a moderator of overall life satisfaction—Life facet satisfaction relationships. *Social Indicators Research*, 40: 259–284.
- Martin, R. L., & Osberg, S. 2007. Social entrepreneurship: The case for definition. *Stanford Social Innovation Review*, 5: 29–39.
- Matten, D., & Moon, J. 2008. “Implicit” and “explicit” CSR: A conceptual framework for a comparative understanding of corporate social responsibility. *Academy of Management Review*, 33: 404–424.
- Matthews, J. R. 2011. Assessing organizational effectiveness: The role of performance measures. *The Library Quarterly*, 81: 83–110.
- Mayne, J. 2001. Addressing attribution through contribution analysis: Using performance measures sensibly. *The Canadian Journal of Program Evaluation*, 16: 1–24.
- McLoughlin, J., et al. 2009. A strategic approach to social impact measurement of social enterprises. The SIMPLE methodology. *Social Enterprise Journal*, 5: 154–178.
- McMullen, J. S. 2010. Delineating the domain of development entrepreneurship: A market-based approach to facilitating inclusive economic growth. *Entrepreneurship Theory and Practice*, 35: 185–193.
- Meyer, B. D. 1995. Natural and quasi-experiments in economics. *Journal of Business & Economic Statistics*, 13: 151–161.
- Miller, T. L., Grimes, M. G., McMullen, J. S., & Vogus, T. J. 2012. Venturing for others with heart and head: How compassion encourages social entrepreneurship. *Academy of Management Review*, 37: 616–640.
- Miller, T. L., & Wesley, C. L. 2010. Assessing mission and resources for social change: An organizational identity perspective on social venture capitalists’ decision criteria. *Entrepreneurship Theory and Practice*, 34: 705–733.
2008. In A. Nicholls (Ed.), *Social entrepreneurship: New models of sustainable social change*. Oxford: Oxford University Press.
- Nicholls, A. 2009. “We do good things, don’t we?”: “Blended value accounting” in social entrepreneurship. *Accounting, Organizations and Society*, 34: 755–769.
- Nicholls, A., & Cho, A. 2008. Social entrepreneurship: The structuration of a field. In A. Nicholls (Ed.), *Social entrepreneurship: New models of sustainable social change*: 99–118. Oxford: Oxford University Press.
- OECD. 2014. *Society at a glance 2014: OECD social indicators*. http://dx.doi.org/10.1787/soc_glance-2014-en.
- Oswald, A. J., & Wu, S. 2010. Objective confirmation of subjective measures of human well-being: Evidence from the U.S.A. *Science*, 327: 576–579.
- Parsons, T. 1956. Suggestions for a sociological approach to a theory of organizations. *Administrative Science Quarterly*, 1: 63–85.
- Paton, R. 2003. *Managing and measuring social enterprises*. London: Sage.
- Petrucchi, A., & D’Andrea, S. S. 2002. Quality of life in Europe: Objective and subjective indicators: A spartial analysis using classification techniques. *Social Indicators Research*, 60: 55–88.
- Phills, J. A., Deiglmeier, K., & Miller, D. T. 2008. Rediscovering social innovation. *Stanford Social Innovation Review*, 6: 34–43.
- Polonsky, M., & Grau, S. L. 2011. Assessing the social impact of charitable organizations—Four alternative approaches. *International Journal of Nonprofit and Voluntary Sector Marketing*, 16: 195–211.
- Price, J. L. 1972. The study of organizational effectiveness. *The Sociological Quarterly*, 13: 3–15.
- Quinn, R. E., & Rohrbaugh, J. 1983. A spatial model of effectiveness criteria: Towards a competing values approach to organizational analysis. *Management Science*, 29: 363–377.

- Reed, J., Jones, D., & Irvine, J. 2005. Appreciating impact: Evaluating small voluntary organizations in the United Kingdom. *Voluntas: International Journal of Voluntary and Nonprofit Organizations*, 16: 123–141.
- Reinhardt, C. S. 2011. Evaluating methods for estimating program effects. *The American Journal of Evaluation*, 32: 246–272.
- Rojas, M. 2006. Life satisfaction and satisfaction in domains of life: Is it a simple relationship? *Journal of Happiness Studies*, 7: 467–497.
- Rojas, R. R. 2000. A review of models for measuring organizational effectiveness among for-profit and nonprofit organizations. *Nonprofit Management & Leadership*, 11: 97–104.
- Rossouw, S., & Pacheco, P. 2012. Measuring non-economic quality of life on a sub-national level: A case study of New Zealand. *Journal of Happiness Studies*, 13: 439–454.
- Ryan, P. W., & Lyne, I. 2008. Social enterprise and the measurement of social value: Methodological issues with the calculation and application of the social return on investment. *Education, Knowledge & Economy*, 2: 223–237.
- Sandvik, E., Diener, E., & Seidlitz, L. 1993. Subjective well-being: The convergence and stability of self-report and non-self-report measures. *Journal of Personality*, 61: 317–342.
- Santos, F. M. 2012. A positive theory of social entrepreneurship. *Journal of Business Ethics*, 111: 335–351.
- Sawhill, J. C., & Williamson, D. 2001. Mission impossible: Measuring success in nonprofit organizations. *Nonprofit Management & Leadership*, 11: 371–386.
- Schimmack, U. 2008. *The structure of subjective wellbeing*. Working paper. Department of Psychology, University of Toronto.
- Scholten, P. 2005. *Social return on investment, performance measurement for social entrepreneurs*. Paper presented at the IESE Social Enterprise Conference, Barcelona.
- Schuetz, A. 1967. *Collected papers 1: The problem of social reality*. The Hague: Nijhoff.
- Schyns, P. 1998. Crossnational differences in happiness: Economic and cultural factors explored. *Social Indicators Research*, 43: 3–26.
- Sen, A. 1985. Well-being, agency and freedom: The Dewey lectures 1984. *The Journal of Philosophy*, 82: 169–221.
- Sen, A. 1999. The possibility of social choice. *The American Economic Review*, 89: 349–378.
- Short, J. C., Moss, T. W., & Lumpkin, G. T. 2009. Research in social entrepreneurship: Past contributions and future opportunities. *Strategic Entrepreneurship Journal*, 3: 161–194.
- Sorensen, J. E., & Grove, H. D. 1977. Cost-outcome and cost-effectiveness analysis: Emerging nonprofit performance evaluation techniques. *Accounting flevieu*, 52: 658–675.
- Sowa, J. E., Selden, S. C., & Sandfort, J. R. 2004. No longer unmeasurable? A multidimensional integrated model of nonprofit organizational effectiveness. *Nonprofit and Voluntary Sector Quarterly*, 33: 711–728.
- Stame, N. 2010. What doesn't work? Three failures, many answers. *Evaluation*, 16: 371–387.
- Steers, R. M. 1975. Problems in the measurement of organizational effectiveness. *Administrative Science Quarterly*, 20: 546–558.
- Stiglitz, J., Sen, A., & Fitoussi, J.-P. 2009. *The measurement of economic performance and social progress revisited—Reflections and overview*. Commission on the Measurement of Economic Performance and Social Progress: <http://www.stiglitz-sen-fitoussi.fr/documents/overview-eng.pdf>.
- Suldo, S. M., Savage, J. A., & Mercer, S. H. 2014. Increasing middle school students' life satisfaction: Efficacy of a positive psychology group intervention. *Journal of Happiness Studies*, 15: 19–42.
- Swanson, L. A., & Di Zhang, D. 2010. The social entrepreneurship zone. *Journal of Nonprofit & Public Sector Marketing*, 22: 71–88.
- Tan, W.-L., Williams, J., & Tam, T.-M. 2005. Defining the "social" in "social entrepreneurship": Altruism and entrepreneurship. *The International Entrepreneurship and Management Journal*, 1: 353–365.
- Tay, L., Chan, D., & Diener, E. 2013. The metrics of societal happiness. *Social Indicators Research*, 117: 577–600.
- Tilley, N. 1996. Demonstration, exemplification, duplication and replication in evaluation research. *Evaluation*, 2: 35–50.
- Trelstad, B. 2008. Simple measures for social enterprise. *Innovations: Technology, Governance, Globalization*, 3: 105–118.
- Tsui, A. S. 1990. A role set analysis of managerial reputation. *Organizational Behavior and Human Performance*, 34: 64–94.
- Turker, D. 2008. Measuring corporate social responsibility: A scale development study. *Journal of Business Ethics*, 85: 411–427.
- United Nations. 2013. *Millennium development goals report 2013*. <http://www.un.org/millenniumgoals/pdf/report-2013/mdg-report-2013-english.pdf>.
- Ura, K., Alkire, S., Zangmo, T., & Wangdi, K. 2012. *An extensive analysis of GNH index*. Centre for Bhutan Studies: <http://www.grossnationalhappiness.com>.
- Vanclay, F. 2003. International principles for social impact assessment. *Impact Assessment and Project Appraisal*, 21: 5–11.
- Veenhoven, R. 2000. Wellbeing in the welfare state: Level not higher, distribution not more equitable. *Journal of Comparative Policy Analysis*, 2: 91–125.
- Veenhoven, R. 2009. Well-being in nations and well-being of nations: Is there a conflict between individual and society? *Social Indicators Research*, 91: 5–21.
- Webb, R. J. 1974. Organizational effectiveness and the voluntary organization. *Academy of Management Journal*, 17: 663–667.

- Weerawardena, J., & Mort, G. S. 2006. Investigating social entrepreneurship: A multidimensional model. *Journal of World Business*, 41: 21–35.
- Weiss, C. H. 1997. How can theory-based evaluation make greater headway? *Evaluation Review*, 21: 501–524.
- Whetten, D. A., & Cameron, K. S. 1994. Organizational effectiveness: Old models and new constructs. In J. Greenberg (Ed.), *Organizational behavior: The state of the science*: 135–153. Hillsdale, NJ: Lawrence Erlbaum Associates.
- White, H. 2010. A contribution to current debates in impact evaluation. *Evaluation*, 16: 153–164.
- White, H. 2013. An introduction to the use of randomized control trials to evaluate development interventions. *Journal of Development Effectiveness*, 5: 30–49.
- Woolcock, M. 2013. Using case studies to explore the external validity of “complex” development interventions. *Evaluation*, 19: 229–248.
- Wu, C.-H. 2008. Can we weight satisfaction score with importance ranks across life domains? *Social Indicators Research*, 86: 469–480.
- Yates, B. T. 2009. Cost-inclusive evaluation: A banquet of approaches for including costs, benefits, and cost-effectiveness and cost-benefit analyses in your next evaluation. *Evaluation and Program Planning*, 32: 52–54.
- Yuchtman, E., & Seashore, S. E. 1967. A system resource approach to organizational effectiveness. *American Sociological Review*, 32: 891–903.
- Zahra, S., Gedajlovic, E., Neubaum, D., & Schulman, J. 2009. A typology of social entrepreneurs: Motives, search processes and ethical challenges. *Journal of Business Venturing*, 24: 519–532.
- Zahra, S. A., Rawhouser, H. N., Bhawe, N., Neubaum, D. O., & Hayton, J. C. 2008. Globalization of social entrepreneurship opportunities. *Strategic Entrepreneurship Journal*, 2: 117–131.
- Zammuto, R. F. 1982. *Assessing organizational effectiveness: Systems change, adaption, and strategy*. Albany: State University of New York.
- Zammuto, R. F. 1984. A comparison of multiple constituency models of organizational effectiveness. *Academy of Management Review*, 9: 606–616.

Arne Kroeger (arne.kroeger@ufo.uni-hannover.de) is a doctoral candidate at the Institute of Management and Organizational Behavior at the Leibniz University Hanover and has worked as an investment analyst for the Social Venture Fund, Munich. His current research interests include measuring and comparing social impact, impact investing, and scaling social impact.

Christiana Weber (christiana.weber@ufo.uni-hannover.de) holds the chair for management and organizational behavior at the Leibniz University Hanover. She received her Ph.D. in sociology from the Social Science Research Centre, Berlin (WZB)/Technical University Berlin. Her current research interests include (social) entrepreneurship, corporate sustainability, social network theory, and innovation management.

Copyright of Academy of Management Review is the property of Academy of Management and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.