Creating Flexible Standards:
Construct-based equivalence and bounded flexibility

Executive Summary

Current impact measurement approaches can be grouped into one of two categories: *flexible approaches* that emphasize bespoke indicators based on a learning process within the organization; and *uniform approaches* that seek to create and promote shared indicators.

Research on standard-setting suggests that successful standards – those with wide and long-lasting adoption - must balance uniformity and flexibility. As Timmermans and Epstein (2010) note in their review, "a recurring surprising finding is that loose standards with great adaptability may work better than rigidly defined standards….The trick in standardization appears to be to find a balance between flexibility and rigidity"(p 81).

A successful social impact standard, therefore, will be neither uniform indicators nor bespoke indicators. It must find a middle ground between anything-goes and one-right way.

This report outlines how a flexible standard can be created for social and environmental indicators. It introduces to mechanisms, *construct-based equivalence* and *bounded flexibility*, that are used in Financial Accounting - perhaps the best example of a standard that manages to somehow be both the same but different – to achieve this balancing act between uniformity and flexibility.

The flexible standard is illustrated using an example of 1110 indicators gathered from 114 organization. It shows that using the flexible approach, middle ground can be found between the bespoke process-based approaches for creating indicators, and rigid sets of uniform indicators.

The problem

Current impact measurement approaches can be grouped into one of two categories: *flexible approaches* that emphasize bespoke indicators based on a learning process within the organization; and *uniform approaches* that seek to create and promote shared indicators.

The advantage of bespoke indicators is relevance to context. Indicators that are designed to reflect the context allow for better organizational learning and innovation. Jane Addams was ab early proponent of flexibility in impact measurement through work with Hull House beginning in 1889 (Oakes & Young, 2008). Current examples of these approaches include theory-of-change, outcome mapping, chain of results, impact map, and logic model. Today, most organizations use these flexible approaches.

Proponents of uniform indicators argue that measurement will be more useful if organizations measured the same things in the same ways. They point to the possibilities of greater insights about what works and the cumulative impacts of collaborations. There have been *many* attempts to create *the* list of uniform indicators for a particular field, at least as early as Allen (1906) proposed indicators for hospitals. Recent examples include indicator banks (Wadia and Parkinson 2011) such as Urban Institute's Outcomes Project (2004), United Nation Sustainable Development Goals (UN SDGs), Global Indicator Framework (2017), and collective impact (Kania & Kramer, 2011) initiatives. Newer efforts to create a uniform set of indicators are themselves an indication that earlier attempts failed to take root as *the* standard. Most uniform indicators fail. Literature on standard setting may explain why

Why we need flexible standards

Most people think of standards need to be uniform to be useful, but in fact, research has argued that flexible standards are more informative (Hann, Lu, and Subrananyam 2007) and comparable (Dye and Verracchia 1995; Merino and Coe 1978). They are more widely adopted (Brunsson and Jacobsson 2002; Sandholtz 2002) and they last longer (Egyedi and Blind 2008). It is not just that flexibility is good, it is that total uniformity is harmful. Uniform indicators undermine the organization's autonomy to measures what matters (Baur & Schmitz, 2012). Uniform indicators undermine stakeholder's voice because stakeholders are not able to articulate success in their own terms (Grey et al. 1997; Brown and Dillard 2014) instead relying on some central indicator-setting body to define what is "good" (Tregidga, Milne and Kearins 2018). Moreover, indicators can thwart innovation by requiring organization to conform to someone else's goals (Campbell, 2003) and impede learning by limiting the relevance of information.

How to Create a Flexible Standard

Rather than building consensus around a single set of indicators, those seeking a common approach to impact measurement should identify key constructs and craft the rules and principles of flexibility in measuring those constructs.

Construct equivalence

Constructs are ideas or concepts. Youth, vulnerability, poverty and food insecurity are all constructs. Construct-based equivalence treats items as equivalent if they are conceptually the same, even if they are not measured in the same way across entities. In accounting, inventory *is* inventory whether it measured by FIFO (first-in, first-out) or weighted average. Construct-based equivalence can be juxtaposed with measurement-based equivalence, the prevailing approach to social and impact measurement. Under measurement-based equivalence, standard setters focus on measuring the *exact* same way (e.g. youth *is* ages 15-24) which makes other reasonable measures of the construct (18-30 years) somehow wrong.

Bounded flexibility

Bounded flexibility uses rules and/or principles to put limits on what is considered equivalent. Bounded flexibility narrows the range of options, finding a middle ground between one-right-way and anything-goes (Ruff & Olsen 2016). There are many valid measures of youth, vulnerability, poverty and food insecurity. But not *all* measures are valid (50 years is not youth!). Using bounded flexibility, there can be many allowable alternative *indicators* of a construct. This allows the measurement of the construct to be relevant to many different organizations and constructs, ensuring adaptability in numerous situations. Two mechanisms by which flexibility is bounded are if:then and call-shot.

If:then flexibility

If:then specifies a particular method for a particular operating context. There is only one allowable method for a given context. The task of the standard-setter is then to define the relevant circumstances that should be taken into account.

A Call-shot flexibility approach

Call-shot offers organizations discretion in methods and metrics, so long as the method is disclosed. The term is a reference to a game of pool, in which players have a wide discretion as to which ball to shoot into which pocket but must call the shot ahead of time correctly. This analogy is used to highlight the three key elements of this flexibility: discretion, disclosure, and follow-through.

Using this flexible standard, each organization defines its own indicators — within limits — and construct-equivalence is used to represent those unique flexible indicators as common ideas.

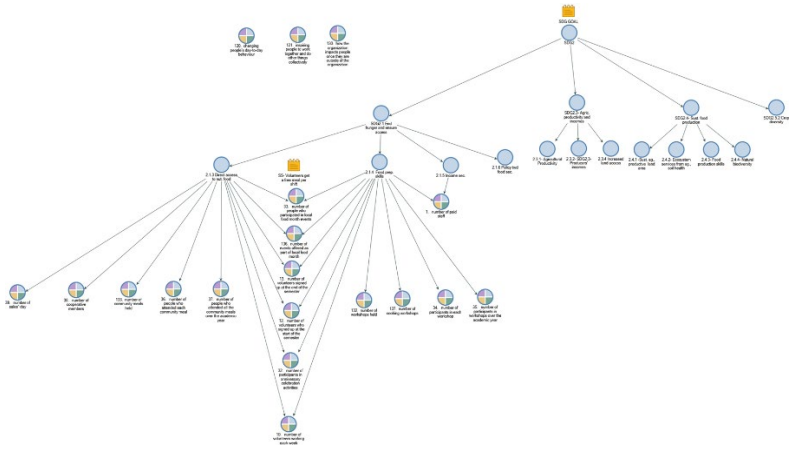
Illustrative example of a flexible standard

To examine how these might work with social impact standards, a study applied construct equivalence and bounded flexibility to a sample of 1110 social impact indicators from 114 organizations (Ruff 2019).

Identify common constructs

In the study all 144 indicators related to hunger and sustainable food systems can be expressed as one of six constructs and all 233 indicators related to employment and access to financial products can be expressed as eight constructs. Figure 1 illustrates the nested relationship of indicators and constructs.

Figure 1: Indicators were grouped into common constructs



Use constructs to tell a common story across organizations

The common constructs were used to tell a common story across the many organizations, without creating a uniform set of indicators. Panel 1 illustrates how three different organization might report on the same set of four ideas related to employment, each using its own specific indicators. This highlights that it is possible to tell a common story across the 114 organizations.

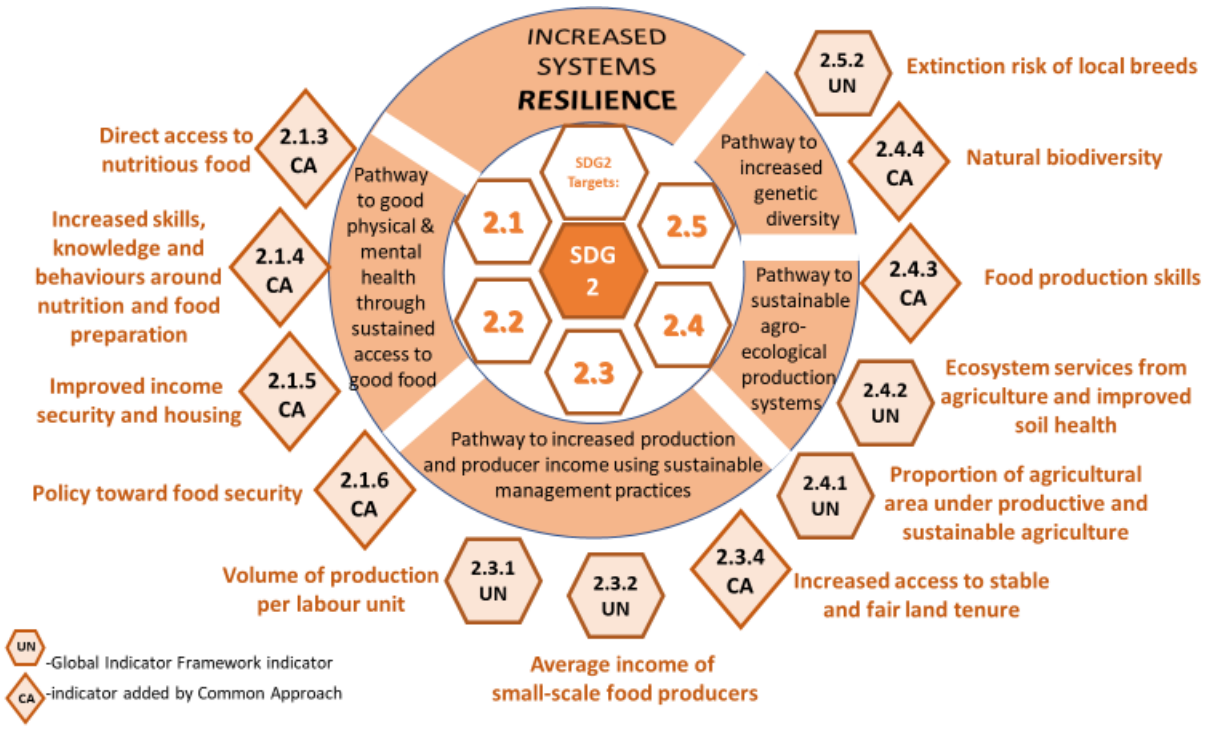
Panel 1: Making things the same with construct-based equivalence: An illustrative example of the headline indicator "People with improved employment situation" using 3 different organizations (Ruff 2019)

1	This organization provides workplace internships to help develop employment skills		
	People progressing on a path to employment		
	Measure: clients who acquire new technical skills	21	
	Measure: clients who say they feel good about themselves	18	
	Measure: clients with increased awareness of money and the economy	15	
	Measure: clients with improved self-care	12	
	Notes: All assessed by employment counsellor; duplicates removed; in total		23
	People acquiring a good job		
	Measure: Clients who find full-time, minimum-wage employment or better during the period		3
	People retaining a good job		
	Measure: Clients still employed after 12-week follow-up		0
	People with improved employment situation		**30**
	Good jobs provided directly by organization		
	Measure: full-time pay with benefits and compensation at 50th percentile or higher		
	Notes: These are staff positions		**4**
2	This organization provides dignity-through-work and social interaction for adults with significant cognitive impairments.		
	People progressing on a path to employment		N/A
	People acquiring a good job		
	Measure: New hires during the period		2
	Notes: Administrative records		
	People retaining a good job		41
	A good job is one that has working conditions respond to the needs and capacities of each targeted employee; has opportunities to develop skills and knowledge; enables employees to contribute to the best of their abilities and provides stimulating work areas. Includes new hires.		
	Good jobs provided directly by organization		**44**
	Measure: These are staff positions plus social hires		
	Notes: See policy on our definition of a good job for staff positions.		
3	This organization is a media firm and certified B Corp		
	Good jobs provided directly by organization		**68**
	Measure: Based on responses to the ten B Assessment Survey questions pertaining to quality of employment. If 80% or more agree or strongly agree on each question we assume 80% of jobs are "good."		

Link to the UN Sustainable Development Goal (SDGs) Global Indicator Framework

Further, the study compared the 1110 indicators to the UN Sustainable Development Goal (SDGs) Global Indicator Framework. The study finds that construct-based equivalence and bounded flexibility can improve the entity-level relevance of the standard. In fact, without construct-based equivalence and bounded flexibility only 24% of enterprise-defined indicators match or somewhat match to the SDGs. However, with construct-based equivalence and bounded flexibility 100% of indicators match. Figure 2 shows how the Common Approach created constructs link with the UN SDG targets and goals using the example of indicators related to SDG 2: End hunger, achieve food security and improved nutrition and promote sustainable agriculture

Figure 2: constructs can link diverse indicators to SDG targets and goals



Limitations and considerations

Marginalized and vulnerable voices

Social impact standards create new types of knowledge. Such standards could amplify and lend visibility to the voices of people, future generations, and the ecosystem in decision-making venues where their interests have been historically overlooked. However, historically standards have often been a means by which powerful groups have enacted the world to their benefit. A flexible standard will be better than a uniform one at accommodating a diversity of perspectives. However, even the process of articulating common constructs and the rules of bounded flexibility, it will be very important to attend to *who's* views are being represented.

Shifting meanings

The fact that different people define "good" differently highlights a challenge with all impact standards, one that is not remedied by construct-equivalence or bounded flexibility. Defining the constructs themselves is a social process entangled with moral and ethical positions. As such, the same measures can mean different things for different organizations. The moral complexity arises in the links between the counts and the constructs they are meant to indicate. The relations are non-linear, dynamic, and intrinsically entangled with moral values. Thus, while the countable units can be objective and auditable, the relations that organizations draw between indicators and constructs is often one of hunch or lived experience, which report readers don't have access to.

Key Recommendations

Focus on bounded flexibility

Standard-setting should focus on crafting the rules and principles of bounded flexibility, rather than on agreement of uniform indicators. Organizations can be granted authority to select their own indicators in dialogue with stakeholders, as recommended by the Global Reporting Initiative and UN Global Compact (2018). A flexible standard has more relevance, is more informative, comparable, widely adopted, longer-lived, and relevant for diverse users.

Flexibility is created by:
1) Articulating common constructs.
2) Giving organizations wide discretion on how they measure those constructs.

Favouring construct-equivalence to measure-equivalence

Forgoing measure-equivalence in favour of construct-equivalence opens a flexible standard to polyvocal perspectives, which would bring social impact standards more in line with the stakeholder dialogue practices widely considered to be best practice in the field of impact evaluation. The ability to be both the-same-and-different appears to solve many current problems with impact measurement standards. This has implications for the adoption and longevity of an impact standard. It also has ethical implications, as more perspectives can be voiced through said standard.

Use of both call-shot and if:then flexibility approaches

Both if:then and call-shot approaches may be employed to facilitate such a standard, ensuring polyvocal representations and high adoption of the standard. The articulations of these limits allows standard-setters to dial up or dial down the balance between uniformity and flexibility.

Resources Consulted

Allen, W. H. (1906). Hospital Efficiency. *American Journal of Sociology*, 12(3), 298–318.

Baur, D., & Schmitz, H. P. (2011). Corporations and NGOs: When Accountability Leads to Co-optation. *Journal of Business Ethics*, 106(1), 9–21.

Brown, J., & Dillard, J. (2014). Integrated reporting: On the need for broadening out and opening up. *Accountability, Auditing & Accountability Journal*, 27(7), 1120–1156.

Brunsson, N., & Jacobsson, B. (2002). *A World of Standards*. Oxford, UK: Oxford University Press, 198.

Campbell, D. (2003). Outcomes Assessment and the Paradox of Nonprofit Accountability. *Nonprofit Management and Leadership*, 12(3), 243–259.

Cole, V., Branson, J., & Breesch, D. (2012). The uniformity-flexibility dilemma when comparing financial statements: Views of auditors, analysts and other users. *International Journal of Accounting and Information Management*, *20*(2), 114–141.

Dye, R. A., & Verrecchia, R. E. (1995). Discretion vs. Uniformity: Choices Among GAAP. *Accounting Review*, *70*(3), 389–415.

Egyedi, T. M., & Blind, K. (2008). *The Dynamics of Standards*. Cheltenham, UK: Edward Elgar Publishing.

Global Reporting Initiative (GRI) & United Nations Global Compact. (2018). *Integrating the SDGs into Corporate Reporting: A practical guide.* https://www.globalreporting.org/resourcelibrary/GRI_UNGC_Reporting-on-SDGs_Practical_Guide.pdf (Retrieved 04/20/2019)

Gray, R., Dey, C., Owen, D. L., Evans, R., & Zadek, S. (1997). Struggling with the praxis of social accounting: stakeholders, accountability, audits and procedures. *Accounting, Auditing & Accountability Journal*, *10*(3), 325–364.

Hann, R. N., Lu, Y. Y., & Subramanyam, K. R. (2007). Uniformity versus Flexibility: Evidence from Pricing of the Pension Obligation Author. *The Accounting Review*, *82*(1), 107–137.

Kania, J. & Kramer, M. (2011). Collective Impact. *Stanford Social Innovation Review*.

Merino, B. D., & Coe, T. L. (1978). Uniformity in Accounting: An illustrative overview of uniformity - from generations ago to the Metcalf and Moss reports. *The Journal of Accountancy*, (August), 83–91.

Oakes, L. S., & Young, J. J. (2008). Accountability re-examined: Evidence from Hull House. *Accounting, Auditing and Accountability Journal*, *21*(6), 765–790.

Ruff, K. (2019). Good Food and Good Jobs: How Construct-based Equivalence and Bounded Flexibility can increase entity-level relevance of social and environmental accounting standards. https://papers.ssrn.com/sol3/cf_dev/AbsByAuth.cfm?per_id=2904156 (Retrieved 04/24/19).

Ruff, K. & Olsen, S. (2016). The Next Frontier in Impact Measurement isn't Measurement at All: Why We Need Skilled Analysts to Improve Social Capital Markets. *Stanford Social Innovation Review*, May 10. ssir.org/articles/entry/the_next_frontier_in_social_impact_measurement_isnt_measurement_at_all Retrieved (04/17/19).

Sandholtz, K. (2012). Making Standards Stick: A Theory of Coupled vs. Decoupled Compliance. *Organization Studies* 33 (5–6): 655–79.

Timmermans, S. & Epstein, S. (2010). A World of Standards but not a Standard World: Toward a Sociology of Standards and Standardization. *Annual Review of Sociology* 36 (1): 69–89. doi.org/10.1146/annurev.soc.012809.102629.

Tregidga, H., Milne, M., & Kearins, K. (2014). (Re)presenting "sustainable organizations." Accounting, Organizations and Society, 39(6), 477–494.

United Nations. (2018). Global Indicator Framework for the Sustainable Development Goals and Targets of the 2030 Agenda for Sustainable Development E/CN.3/2018/2, Annex II. https://unstats.un.org/sdgs/indicators/indicators-list/.

Wadia, A., & Parkinson, D. (2011). Outcome and Outcome Indicator Banks: Availability and Use, Charities Evaluation Services. https://www.globalgiving.org/social-impact-academy-static/pdf/ces_outcomes_and_outcome_indicator_banks.pdf