The Use of Weightings in Context-Based Sustainability and the MultiCapital Scorecard – A Brief Introduction

Prepared by Mark W. McElroy, PhD
Center for Sustainable Organizations

Context-Based Sustainability (CBS) and the MultiCapital Scorecard (MCS) are measurement methods and tools that can be used to assess the performance of organizations and other human social systems. The MCS, in particular, takes the form of an index (i.e., scorecard) composed of multiple areas of impact (or resource types and categories), each of which and its corresponding metric(s) serves as an indicator of organizational performance.

In the MCS, the general theory of performance we rely on is, in fact, *sustainability* performance – and not just for social and environmental impacts. In other words, the standard by which we judge the performance of a collective is the degree to which it is operating *sustainably* on all fronts: social, environmental and economic. This accounts for the fact that in the MCS, we generally organize our indicators in terms of the so-called *triple bottom line*, and also for why we assess performance in terms of impacts on the sufficiency of vital resources (i.e., capitals).

In what follows below, we explain the role that weightings can play in the use of CBS/MCS in light of the explanation above. In total, there are two forms of weighting used in CBS/MCS: 1) weightings of importance, and 2) weightings of indicator-strength.

Importance Weightings

Not all dimensions of performance in an organization will be of equal importance from the perspective of its managers and directors. Indeed, for some areas of impact, poor performance may be more acceptable than for others, in which case their influence on overall scores should be downsized accordingly. The opposite is true as well.

With these variable degrees of importance in mind, CBS and MCS make it possible to assign different weights of importance to different areas of impact, using a simple rating scale. While users of CBS/MCS are free to employ their own such scales, the most common importance scale used is 1–5, where 1 = lowest priority and 5 = highest priority.[1] Users of CBS/MCS are otherwise free to employ importance scales of whatever sort they like or even to refrain from using importance weightings at all. CBS/MCS can be used with or without them.

When importance weights are in fact used, we recommend they be assigned from a fixed budget of available points so that the relative priority of each area relative to all the others can be more rigorously indicated (i.e., weighting points are limited in supply and so must be carefully allocated). The use of a fixed budget of weights, that is, forces organizations to make

[1] See, for example, cases given in Thomas, M. and McElroy, M. 2016. *The MultiCapital Scorecard—Rethinking Organizational Performance*. White River Junction: Chelsea Green Publishing.

hard choices about such relative weights as opposed to, say, caving into the temptation to weight everything at the same high level of importance.

In cases where a scale of, for example, 1–5 is being used, we suggest setting a budget of available points by multiplying the middle value of 3 by the total number of indicators or metrics in use. Thus, an MCS scorecard populated with 20 indicators would produce a budget of 60 points for importance weighting. If the scale is 1–3 (having a middle value of 2) with 9 indicators as shown in Figure 1, the budget would be 18 points. Other ways of establishing a fixed budget could certainly be used, as could the use of scores that are more granular than just whole numbers (i.e., by using scores denominated in tenths or even hundredths of a point for each indicator).

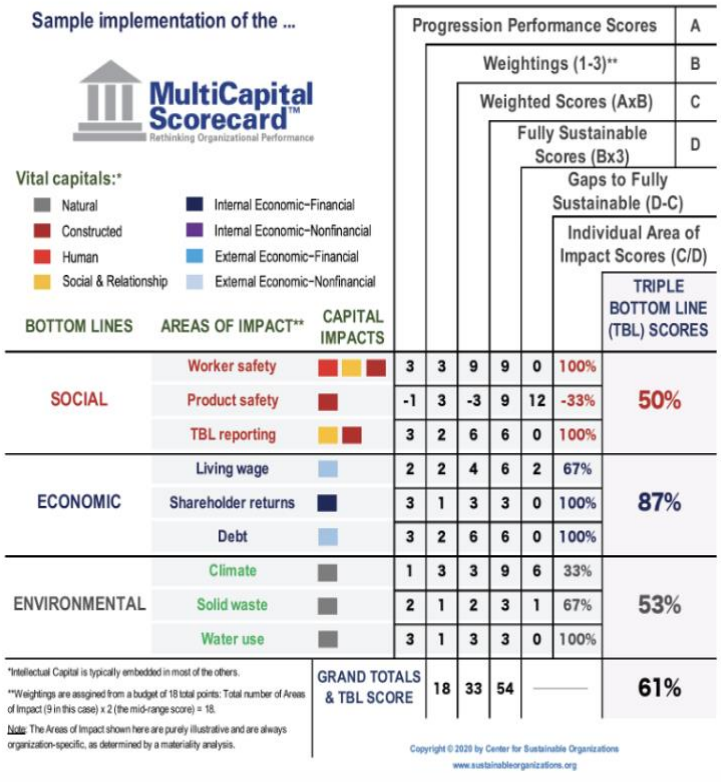

Figure 1 – Importance Weights in a MultiCapital Scorecard

Indicator-Strength Weightings

Because the general theory of performance utilized in CBS/MCS is sustainability per se (i.e., whereby performance is assessed in terms of an organization's impacts on the sufficiency of vital capitals), the strengths of individual indicators and the metrics they entail must also be judged accordingly – that is, by reference to how good a job we think they do in *indicating* or standing for the sustainability performance of organizations. As the well-known social scientist, Paul Lazarsfeld, once put it in a classic 1958 paper of his:

> "… each indicator has not an absolute but only a probability relation to our underlying concept [i.e., to *sustainability performance*, in the case of CBS and the MCS] …"[2]

The social sciences, in turn, are full of methodologies for assigning different weights or strengths (or *probability weights*) to different indicators and their metrics, including, for example, the Analytic Hierarchy Process, or AHP.[3] Most of these methods are fairly complex and expensive to use, in which case more expedient approaches can be applied instead. Such

[2] Lazarsfeld, P. 1958. "Evidence and Inference in Social Research." *Daedalus* 87 (4): 99-130.
[3] Saaty, T. and Vargas, L. 2001. *Models, Methods, Concepts & Applications of the Analytic Hierarchy Process*. Boston: Kluwer Academic Publishers.

alternatives can include, for example, expert-based attempts to subjectively judge the varying strengths of indicators on a simple scale of, say, 0.1 to 1.0, where 0.1 would represent a relatively weak indicator and 1.0 would represent a very strong or perfect one.

In practice, there will rarely be indicators worthy of receiving a strength of 1.0 in the social sciences, since most of what we are trying to measure in management, for example, is too abstract and defies direct measurement. The best we can do is measure other things in their stead, but only when we believe they are reliable proxies or surrogates (i.e., indicators) for what we are really trying to assess – like sustainability performance.

That all said, some indicators are clearly more representative than others when it comes to correlating with or standing for the things we are trying to assess. When attempting to measure the sustainability performance of an organization, for example, the strength of an indicator will vary depending on how far it and the thing it corresponds to are from each other in a causal chain.

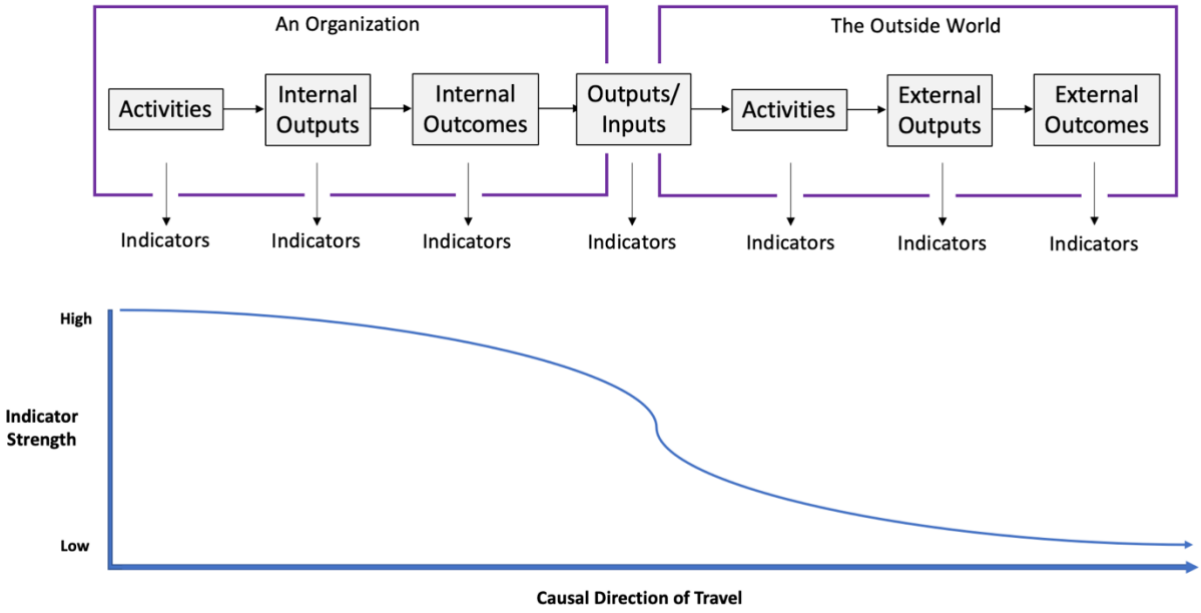

Figure 2 - Causal Chains, Indicators and Indicator-Strength

This phenomenon is illustrated in Figure 2, where the strength of indicators generally declines as one moves from left to right in the diagram, when viewed from the perspective of an organization (i.e., on the left). In other words, the more one relies on indicators on the right side of the diagram to assess the performance of agents on the left, the less reliable they are likely to be.

For the same reasons, organizations will also tend to be held *less accountable* for activities, outputs, or outcomes that are outside of their direct control (i.e., to their right in the diagram), and in turn will be *more accountable* for things that are within their control (i.e., closer to themselves on the left). Indicators that correspond to more causally proximate things,

therefore, can be seen as having greater strength when it comes to signifying and assessing an organization's own performance.

In fact, indicators associated with downstream outcomes that organizations individually are only remotely responsible for, such as climate change, may not be used at all as proxy measures of their performance. Instead, they might only be held accountable for climate-related factors that are more upstream and closer to home, so to speak, if not directly within their control.

This is why we can, and do, hold organizations solely accountable for their own greenhouse gas emissions but not the downstream climate change their emissions bring about – not exclusively, that is. If the thinking behind the causal chain is valid (i.e., if GHG emissions really do cause climate change), then holding organizations accountable for their own emissions is certainly good enough, as are the indicators they use *strong enough*, since what's being measured in the causal chain are their own direct outputs, not the external outcomes way downstream that they each only minimally contribute to.

To consider another example, municipal governments will often have accountability for outcomes in the communities they serve despite the causal distance between anything they might do and results in the field, so to speak. What might be a very important indicator pertaining to, say, the urgent need for affordable housing in a community, might therefore also be a very weak one from a performance assessment standpoint. Indeed, a government's options for action might be limited to low-impact things like how easy or not they can make it for builders to obtain permits.

Thus, if the performance of a local government is being assessed in terms of the effects it might have on affordable housing outcomes, simplifying the issuance of permits will only get them so far in terms of what needs to be done. Indeed, streamlining the issuance of permits to builders is just the tip of the iceberg when it comes to what it might actually take to get affordable housing built to sufficient levels. Most of the other things needed are arguably out of government's control. Thus, impacts on affordable housing could be of the highest priority, even though the strength of an indicator being used to assess performance against it (e.g., how long it takes to process building permits) could be disappointingly low.

As always, though, whether or not to use indicator-strength weights in performance accounting is up to the practitioner, bearing in mind that the time, complexity, and cost of doing so may be high.