

# Measuring progress and well-being: A comparative review of indicators

Chris Barrington-Leigh\* and Alice Escande

Published in *Social Indicators Research*,  
[doi:10.1007/s11205-016-1505-0](https://doi.org/10.1007/s11205-016-1505-0), 2017.

## Abstract

We provide a new database sampling well-being and progress indicators implemented since the 1970s at all geographic scales. Starting from an empirical assessment, we describe and quantify trends in the institutional basis, methodology, and content of indicators which are intended to capture the broadest conceptions of human social progress. We pay special attention to the roles of sustainability and subjective well-being in these efforts, and find that certain types of indicators are more successful in terms of transparency, accountability, as well as longevity. Our taxonomy encompasses money-denominated accounts of “progress”, unaggregated collections of indicators, indices, and measures oriented around subjective well-being. We find that a most promising innovation is the indices whose weights are accountable to empirical data, in particular through models of subjective well-being. We conclude by amplifying others’ advocacy for the appropriate separation of current well-being from environmental indicators, and for the avoidance of aggregation except where it is meaningful.

KEYWORDS: well-being, progress, quality of life, subjective well-being, life satisfaction, sustainable development, genuine progress

---

\*To whom correspondence should be addressed. We are grateful to Lorrie Herbault, Katie Keys, and Julianne Skarha for excellent research assistance; to Michael Abramson, Stefan Bergheim, Jon Hall, John Helliwell, and Raynald Létourneau for helpful discussions; and for funding from the Social Science and Humanities Research Council of Canada and the Fonds de recherche du Québec — Société et culture.

## Contents

<b>1 Trends and patterns</b>	<b>6</b>
1.1 Trends in names . . . . .	6
1.2 Geographic scales and government involvement . . . . .	9
<b>2 A taxonomy of aggregation methods</b>	<b>12</b>
2.1 Systems of money-denominated accounts . . . . .	15
2.2 Collections of indicators . . . . .	18
2.3 Indices . . . . .	20
2.4 Measures of subjective well-being . . . . .	22
2.5 Blurred boundaries . . . . .	24
<b>3 Selection of Indicators</b>	<b>25</b>
3.1 Top-down or theoretical . . . . .	25
3.2 Bottom-up: democratic or empirical . . . . .	27
<b>4 Roles for subjective well-being and sustainability</b>	<b>32</b>
4.1 Subjective reports of well-being . . . . .	32
4.2 Environmental sustainability . . . . .	35
<b>5 Discussion</b>	<b>37</b>
<b>6 Conclusion</b>	<b>42</b>
<b>References</b>	<b>44</b>

## List of Figures

1	Usage of progress and well-being terminology over time . . .	7
2	Geographic scales and indicator longevity . . . . .	10
3	Government and non-government designers . . . . .	11
4	Indicators, Indices, Accounts, and Subjective measures . . . .	14
5	Incorporation of subjective responses into measures of progress	34
6	Incorporation of environmental indicators into measures of progress . . . . .	37

## List of Tables

1	Weights for OECD domains, inferred from a model of life satisfaction . . . . .	31
---	---	----

## Introduction

It has been eight years since the French Presidency commissioned a report to review measures of economic performance and social progress. The report concluded then that “the time is ripe for our measurement system to shift emphasis from measuring economic production to measuring people’s well-being” (Stiglitz, Sen, and Fitoussi, 2009, p.12).

This represents a profound shift, reflecting an increased availability of new statistical measures, new understanding of human experienced well-being, the widespread growth of inequalities that are not revealed by traditional measures of economic performance, an increased public appetite and acceptance of statistical information, and an increased recognition that environmental degradation threatens the predictability of future welfare.

Since that report, another long-awaited shift in the consensus around development and progress has come about in the form of the Sustainable Development Goals (SDGs). In the context of objectives which include sustainable consumption and production, conservation of other life, as well as equity and life quality, all countries are now developing countries (Kroll, 2015). While the SDGs do not specifically address human well-being in the way that the Stiglitz-Sen-Fitoussi Commission did, they represent the increasingly common paradigm of broadening goals for human outcomes and of combining them with measures of sustainability.

What do the world’s statistical figures of merit look like now, in light of the call to embrace more relevant indicators of social progress?

In recent years, several reviews have already described and advocated for desirable properties of well-being indicators based on statistical validity and normative theories about well-being (e.g., Hagerty et al., 2001; Michalos et al., 2011; Veenhoven, 1996; Stiglitz, Sen, and Fitoussi, 2009; Fleurbaey and Blanchet, 2013). We have a complementary aim, namely to survey the breadth of measures that have been implemented, in order to understand trends, commonalities, and remaining distinctions among approaches.

Hicks, (2012) and Land and Michalos, (2016) describe how a great deal of optimism and interest arose in the late 1960s and early 1970s in forging new systems of social accounts which gave integrating frameworks for the measurement of social progress. These initiatives stemmed from the incapacity of our current economic accounting system, with its emphasis on GDP, to adequately measure societal progress. However, according to Hicks, (2012), this golden age of social accounting and the social indicators movement faded for several decades. Its ambitions in regards to the potential for computer technology were possibly optimistic or at least precocious, and it

failed to generate a compelling consensus about a unifying framework.

That nascent “social indicators movement” of the sixties may have germinated properly only recently. The last decades have witnessed a surge in empirical research concerned with notions of social progress, “green accounting”, sustainability, quality of life, and well-being.

Buttressing numerous attempts to build more multidimensional measures of well-being are new data which give quantitative support to the idea that experienced life quality relates to one’s social and physical environment, personal activities, security, political voice and so on, in addition to health, education, and material wealth and income.

These findings have spurred the way to new measures of growth and progress, not just based on economic factors but including a more comprehensive set of social, psychological, and environmental indicators. These “social indicators,” “augmented GDP” and other well being measures have emerged from a desire to measure *real* progress (Australian Bureau of Statistics, 2014; Central Statistics Office, 2012; Porter, Stern, and Green, 2014); to achieve development (United Nations Development Programme, 2015) or sustainable development (Redefining Progress, 2014; Department for Environment Food & Rural Affairs, 2013; UNEP, 2015); to check vitality and well-being of communities, (Community Foundations of Canada, 2013), regions and nations (The Canadian Index of Wellbeing, 2014a) and compare them against each other (OECD Better Life Initiative, 2014); to guide public policy, inform policy makers and so on.

By searching the academic literature, and government and organizational web sites, we have assembled a sample of efforts to measure progress at different geographic scales over the last 40 years. We analyze the resulting database of 82 indicators as follows. [Section 1](#) quantifies some trends related to terminology, geography, and institutions. [Section 2](#) introduces a conceptual classification of indicators and [Section 3](#) considers the observed methods of choosing their content. [Section 4](#) treats the specific content of subjectively measured well-being and environmental conditions, which are both increasingly prominent in conceptions of progress. [Sections 5 and 6](#) synthesize our findings and conclude. Our recommendation is that subjective well-being may play a central role in measures of human outcomes, but that in order to do them justice most long-term environmental indicators must be separated from those focused on current human well-being.

# 1 Trends and patterns

We begin with an account of some broad patterns evident in our database of 82 measures of progress.<sup>1</sup> Our database is intended to be a selected sample, rather than being either exhaustive or representative. However, we believe we have captured many of the prominent efforts at a range of geographic scales and with a somewhat representative selection of formats, themes, and approaches.

The grey bars in the lower panel of [Figure 1](#) show the distribution of start dates of the measures we analyzed. We found 6 measures originating in the 1980s, 16 in the 1990s, 37 in the 2000s, and 17 so far in the 2010s. Below we assess some broad trends evident over this period.

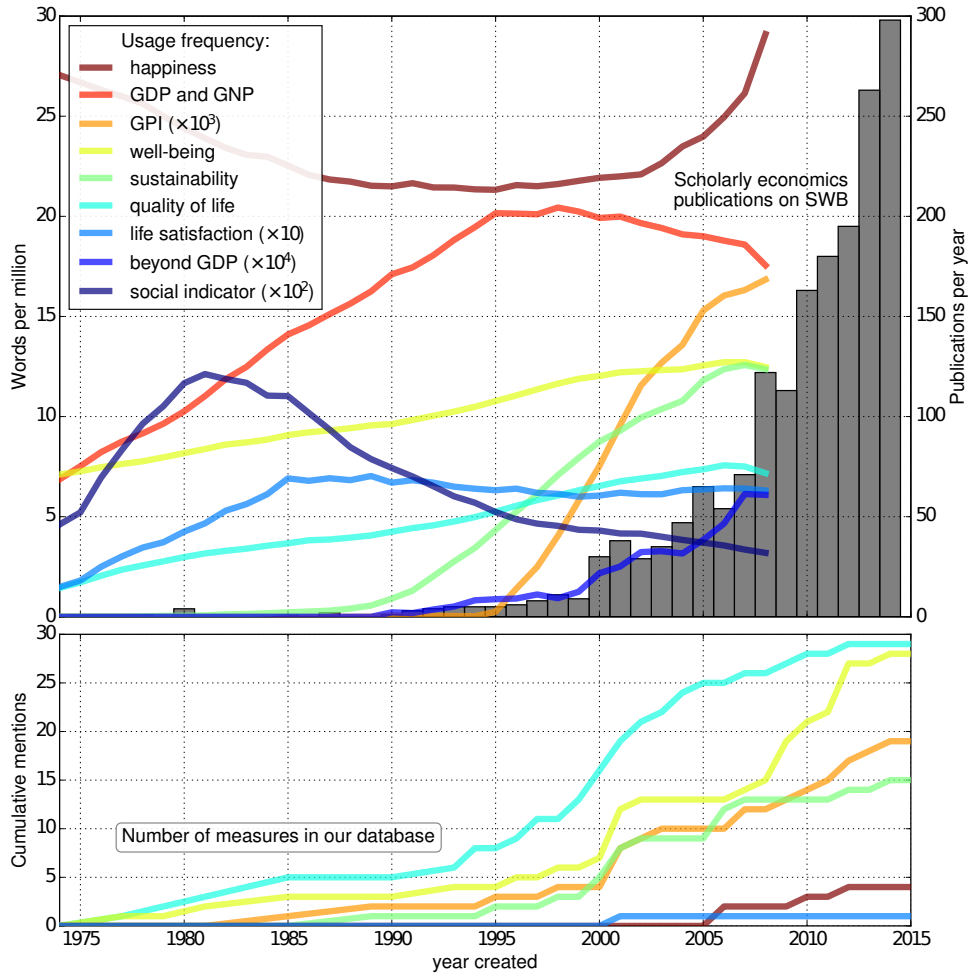
## 1.1 Trends in names

Our database includes indicators with a number of conceptually different rationales and corresponding names, spanning concepts of economic development, generalized wealth, life quality, social development, progress, happiness, and sustainability. Three trends may be discerned in the use of these terms to explain the main function or rationale of each measure. While “quality of life” is a common concept in the content of both recent and earlier measures, the actual *titles* of new measures reflect a general progression from terms like quality of life, “life situation” (The Netherlands Institute for Social Research, 2015), and “social weather” (Social Weather Stations, 2015) towards words related to psychological well-being, including “happiness” and “satisfaction,” along with “well-being” itself. In addition, the word “progress” appears in the names of indicators only after 1995. Lastly, explicit mention of sustainability becomes increasingly common after 2000. In our later discussion, we often use “measure” and “indicator” interchangeably, and use “well-being” and “progress” in their most general senses, in order to encompass the full range of metrics in our database.

Certain patterns in the incidence of words and phrases in the corpus of English books, compiled by Google (see <http://books.google.com/ngrams/>), mirror these trends. As shown by the lines in the upper panel of [Figure 1](#), use of the term “social indicator” peaked around 1980, and the term “genuine progress indicator” appears to have been coined in the mid-1990s. By contrast, “well-being” and “quality of life” are still growing in usage in the English language, and “happiness” is experiencing a resur-

---

<sup>1</sup>This is available online at <http://wellbeing.research.mcgill.ca/publications/WB-indicator-database-2017>.



**Figure 1: Usage of progress and well-being terminology over time.** *Upper panel:* Historical incidences of some relevant terms in printed books, taken from Google’s n-grams. “Life satisfaction” represents the sum of incidences of “life satisfaction” and “satisfaction with life” and is scaled up by a factor of 10 for better visibility. “Social indicator” is scaled up by a factor of 100, and GPI, short for “genuine progress indicator,” is scaled up by 1000. Use of the term “sustainable development” shows a similar pattern over time as “sustainability.” The Google N-gram database ends in 2008. Grey bars show the number of academic economics publications on SWB over time. *Lower panel:* Start dates of well-being / progress measures in our (non-representative) sample (grey bars), along with the (rescaled) cumulative number of mentions of different terms (labeled by colour in top panel) in the stated name or purpose of those measures.

gence of usage since the end of the twentieth century. While the use of “gross domestic product” (GDP) is still on the rise, this is due in recent years to it replacing “gross national product” (GNP), which used to be the favoured measure. The summed incidence of GDP and GNP peaked in the 1990s. By contrast, modern augmented GDP measures, often referred to as “genuine progress indicators” (GPIs) are on the rise, as are “beyond GDP,” “well-being,” “happiness,” and “sustainability.”

We consider some of these same terms when examining the indicators in our database. For each phrase indicated by a coloured line in the lower panel of [Figure 1](#), we track the number of indicators using that phrase in their name or in stating their primary purpose or motivation. Well-being comes out on top, followed by quality of life and progress.

Of course, underlying the observations above are deeper philosophical and conceptual developments, in addition to some technical ones, which relate to the core themes of this paper.<sup>2</sup> One aspect of these developments in popular writing is much starker in the academic literature. Kahneman and Krueger, (2006), when comparing academic studies over the periods 1991–95 and 2001–2005, noted the sharp rise in the number of economics journal articles analyzing data on self-reported life satisfaction or happiness.

The grey bars in the upper panel of [Figure 1](#) show this remarkable trend with one year resolution and continuing to 2015.<sup>3</sup> Although not all new indicators are based on, or incorporate, subjective well-being (SWB), the general growth in interest in more human-based measures of progress may have much to do with the idea that we can now measure happiness quantitatively and with sufficient rigour, along with the specific insights about the importance of social links and context that have been gleaned from this line of research. We devote considerable attention to the role of SWB in [section 4.1](#).

---

<sup>2</sup>These shifts may be said to be not entirely underlain by substance. Land and Michalos, (2016) state that the developers of the Canadian Index of Well-Being regard the term *well-being* to be “roughly synonymous with *overall quality of life*”. Similarly, Thailand and Bhutan both use the word *happiness* in the titles of recent indices which have not much more to do with psychological affect than much earlier analogues (Baramuechai, 2007; Centre for Bhutan Studies, 2015). In addition, even within the narrower context of SWB, the word *happiness* is used as a non-threatening and felicitous informal synonym for subjective well-being, and even life satisfaction in particular, in addition to its narrower specialist meaning as a domain of affect.

<sup>3</sup>These statistics are based on a search for “*life satisfaction*” or “*happiness*” or “*subjective well-being*” in all fields in EconLit, the same economics journal index referenced by Kahneman and Krueger, (2006).



## 1.2 Geographic scales and government involvement

We categorized the indicators in our sample according to their geographic scope. While older indicators in our database tend to have originated as local or national efforts, 37% of the 32 measures initiated since 2005 are international efforts. For instance, the UN’s Inclusive Wealth Index, the OECD’s Better Life Index, Eurostat’s Quality of Life Indicators, and the Social Progress Initiative’s index all arose since 2011 (OECD, 2011; United Nations, 2015; European Statistical System, 2012; *Social Progress Index* 2014). These initiatives come both from international government and from non-government sectors.

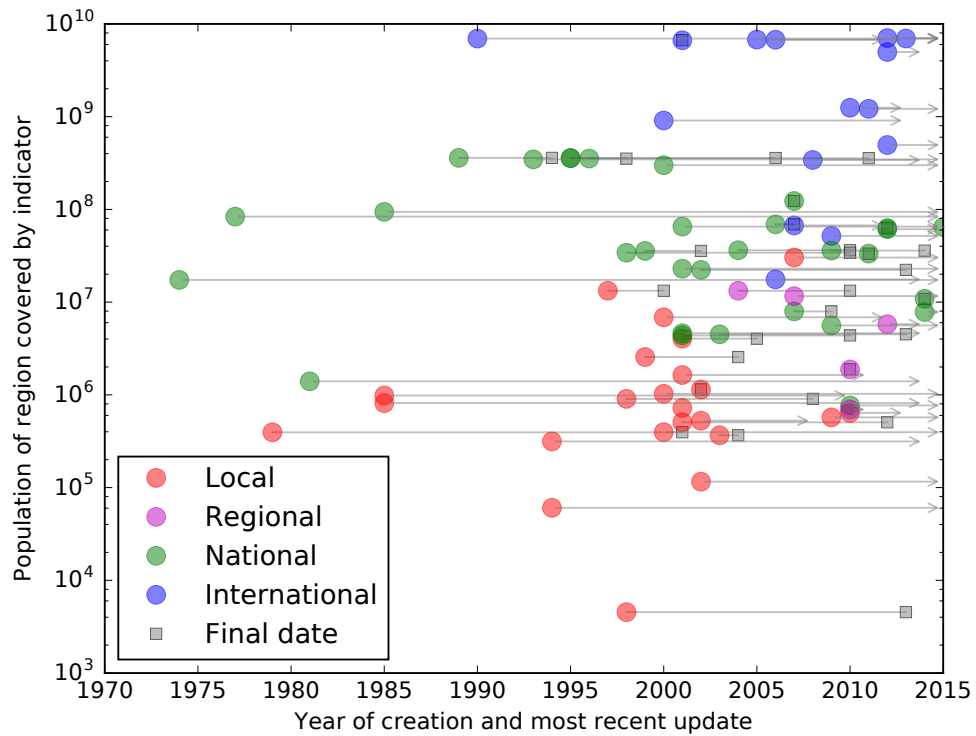
Figure 2 shows the dates, geographic scales, associated populations, and longevity of the initiatives in our database. The regional scale corresponds primarily to provinces and states. Lines ending in arrows mark indicators whose operation we judge to be ongoing, while grey squares show the final years of defunct efforts. The shift in the distribution over time towards larger geographic and population scope is evident, although our earliest indicators were national in origin.

**Local community well-being surveys** Simultaneous with the rise of prominent national and international measures, one reason alternative well-being indicators can be said to be proliferating is because of the rise of local communities’ efforts to build their own accounts of social objectives. Among the oldest of these, located in Jacksonville, Florida, recently released their thirtieth Quality of Life Progress Report (Jacksonville Community Council Inc, 2014). This effort exemplifies a number of features typical of community-organized progress measures.<sup>4</sup> It consists mostly of objective indicators collected from existing statistical data sources; it represents an effort of a community organization or coalition; and it links measured indicators to explicit “goals” and “desired outcomes.”

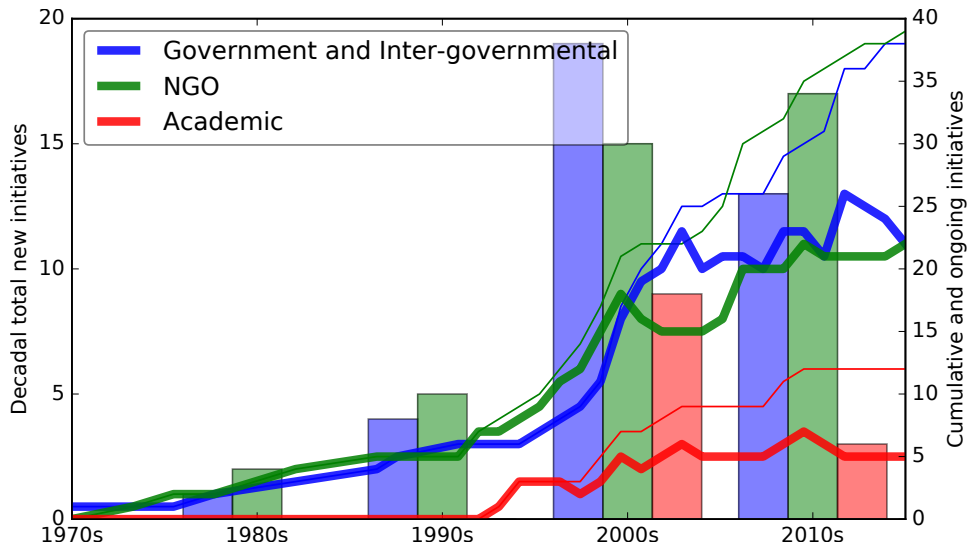
**National governments** By contrast, national-level social indicators tend to be government orchestrated and have arisen particularly in Europe. The measurement of social and economic statistics has historically fallen primarily within the purview of national governments, whether their national statistical agencies or under departmental portfolios. At the national level

---

<sup>4</sup>Especially for local initiatives, our survey database is not intended to be exhaustive. The Jacksonville organization claims its efforts are reflected in over 1000 local community indicators worldwide. One effort to collect links to the growing set of local well-being measurement initiatives is shared at <http://www.communityindicators.net/projects>.



**Figure 2: Geographic scales and indicator longevity.** Start dates for indicators in our database, coloured according to the geographic scale covered by the indicator. The vertical axis shows the population within this region. Gray squares indicate the year in which an indicator was dropped, and arrows show the latest release of the indicator for those which are still in production.



**Figure 3: Government and non-government designers.** Vertical bars show the number of new measures in our database by decade. Two indicators fall simultaneously into the NGO and Academia categories; the rest are in only one. The thin lines give higher time resolution, showing the cumulative number of new measures over time. The thick lines take into account the longevity of indicators, showing the number of surviving measures each year.

also lies, of course, a considerable portion of the policy which might be held accountable to appropriately measured outcomes. Accordingly, with the rise in innovation of new well-being indicators, many initiatives are to be found at the national level. Most of these, in our database, are governmental in origin (The Netherlands Institute for Social Research, 2015; Cooke, 2005; Centres for Disease Control and Prevention, 2014; Department for Environment Food & Rural Affairs, 2013; Perry, 2009; Ministry of Social Development, 2010; Australian Bureau of Statistics, 2014; Central Statistics Office, 2012; Treasury Board of Canada, 2010; Employment and Social Development Canada, 2014; Statistics Finland, 2014; Centre for Bhutan Studies, 2015; Office for National Statistics, 2012).

From the point of view of longevity and impact, there are advantages and disadvantages to governments taking a lead role. Ultimately, influencing the allocation of resources, typically through public policy, is key to the motivation behind any system of indicators. Progress measures which are developed in partnership with governments are the most likely to translate into policy. Moreover, they are more likely to gain traction with a broader

audience including the general public, since other stakeholders with an interest in policy will also invest in understanding and following them.

On the other hand, government-led initiatives may suffer from unreliable financial support or branding and outreach. This is due to an increased vulnerability to changing political mandates, preferences, or identities. For example, the Tasmania Together, Measures of Australia’s Progress, Oregon Benchmarks, First Nations Community Well Being index, the CPRN Quality of Life indicators, the Ontario Quality of Life indicators, and the ESDC Indicators of well-being in Canada were all initiated by governments and have not survived.

Conversely, financial support for the maintenance of an indicator may be more secure through the partnership with a government. While our sample of measures is naturally biased against those which have not survived,<sup>5</sup> of the 33 in our database which have definitively stopped, only 15 were government-designed. [Figure 3](#) compares the number of new and ongoing indicators over time, separated according to whether the effort is led by a government, a private organization, or an individual or group at an academic institution. The latter category is small. The thin lines show the cumulative number of indicators which have been launched, while the thicker lines show the number still running as a function of time. Comparing these two sets of curves shows that while there are similar numbers of government and non-government indicators currently in production, there have been significantly more non-governmental ones launched. That is, the survival rate is higher for government initiatives than for non-governmental ones.

We surmise that due to the advantages and disadvantages we have outlined above, indicators with multi-sectoral coalitions behind them are likely to have the strongest support both initially and over time. This is due in part to the issues we explore next, which may determine the long-run traction enjoyed by different indicator initiatives.

## 2 A taxonomy of aggregation methods

We now address a key issue differentiating the great variety of progress measures. We propose a classification according to how each approach aggregates its various components into an overall measure of progress.

First, it is useful to distinguish between an *index* and a disaggregated

---

<sup>5</sup>That is, our database may underrepresent failed indicator efforts since they no longer have any visibility and were therefore overlooked during our search.

collection of indicators.<sup>6</sup> An index is a scalar value which has been calculated as a weighted sum or other aggregation of some set of constituents. For instance, the U.N.’s Human Development Index is the geometric mean of normalized indices for each of its three component dimensions, one of which is in turn aggregated from its component measures (Human Development Report Office, 2013). Forming an index can capture a concept of progress or well-being in a single value, allowing a summary measure to be tracked over time, plotted easily, and communicated efficiently.

By contrast, some progress measures eschew this sort of aggregation. For instance, the Boston Indicators Project tracks 150 detailed indicators and has organized them into a hierarchy of domains. Like a number of local well-being initiatives, it also organizes small clusters of indicators into goals, and in fact some indicators are even collections of more than one detailed measure. For instance, the number of teachers dedicated to the arts in Boston public schools is a well-defined, scalar indicator which is organized under a goal (“Opportunities for arts education”) which respectively lies within one of ten domains, titled “Cultural Life and the Arts.”

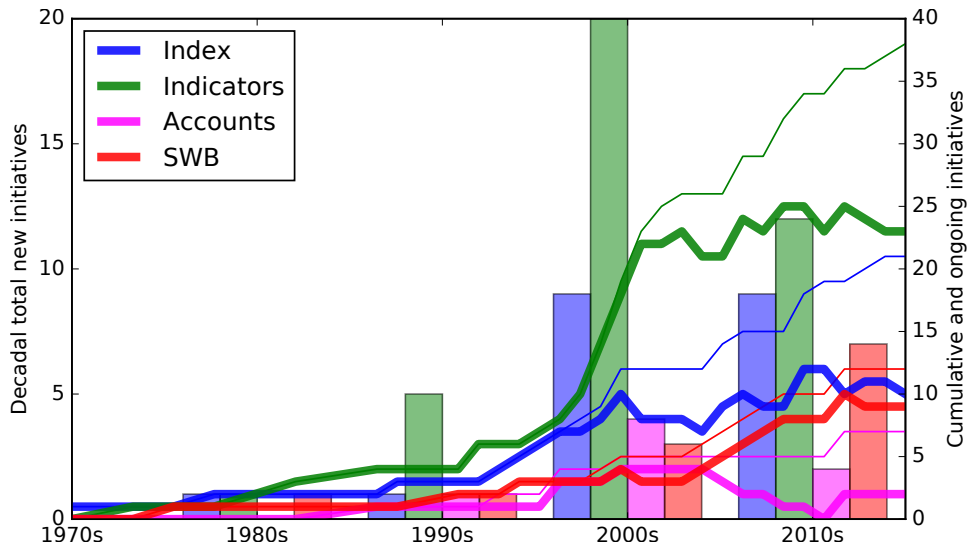
Below we address the advantages and challenges of these approaches, along with a third one, systems of money-denominated accounts, which constitutes a special kind of index,<sup>7</sup> and a fourth, which is when an indicator is based solely on subjective response data. Our database classifies measures (where possible) into accounts, indices, collections of indicators, and measures of subjective well-being. Where aggregation into indices takes place, we note the method, rationale, and weighting involved.

Figure 4 shows the distribution of these types according to their creation date. According to this sample, systems of accounts have a low survival rate from a peak of innovation in the early 2000’s. Indices (excluding those shown under the SWB category) have an overall survival rate of about 50% to date, while unaggregated sets of indicators have fared significantly better. Nearly all of the mostly-recent SWB-oriented indicators are still in production.

---

<sup>6</sup> These are “single indicators” and “component sets” in the language of Shackman, Liu, and Wang, (2005). The second has also been called a dashboard. The terminology can sometimes be awkward, as we use “indicators” and “measures” to include everything, while “sets of indicators,” “unaggregated indicators”, or “dashboards” all refer to one specific category in our database.

<sup>7</sup>Other discussions of some of the advantages and disadvantages of aggregation to indices include Michalos et al., (2011).



**Figure 4: Indicators, Indices, Accounts, and Subjective measures.** Vertical bars show the classification of new measures in our database by decade. The subjective well-being (SWB) category includes measures exclusively composed of subjective assessments, as well as indices aggregated according to weights derived from empirical models of life satisfaction. Excluded from “Index” are those indices which also fit in the SWB category. The thin lines give higher time resolution, showing the cumulative number of new measures over time. The thick lines take into account the longevity of indicators, showing the number of surviving measures each year.

## 2.1 Systems of money-denominated accounts

Many projects of the type included in our database arose in the context of a need to provide an alternative to gross domestic product (GDP) as a measure of social or even economic progress. GDP is itself an index, but it is of a special type because all of its constituents are measured in superficially commensurable dimensions; they are expressed in terms of monetary flows. As a result, weighting issues are less prominent, once components are selected and measured in financial terms. While the history of GDP measures is one of evolving and often political expansion in breadth towards inclusion of previously missing forms of capital and production (Coyle, 2015; Bridgman et al., 2012), one relatively mild critique of the GDP as an overarching social indicator is simply that it still leaves out important components of valuable activity and investment. In response, a movement grew around the calculation of a more complete accounting of investments and production than that which governments were willing to make. Such “augmented GDP” or “genuine progress” measures aim to include the depletion of natural resources, investments in human and even social capital, and forms of production which are valuable but unpaid or which fall outside the formal market. However, including components of provision, investment, and disinvestment or harm — which are not explicitly captured by price-mediated markets — requires higher levels of judgment and extrapolation in order to quantify their contribution. As a result, governments shy away from these activities, even though in principle excluding such factors might require just as much normative justification, from the point of view of the conceptual scope of the GDP.

Seven entries in our database are of the money-denominated account type. Four are named Genuine Progress Indicator (GPI), one is the Inclusive Wealth Index, and the other two refer to economic welfare and economic well-being (Cobb, 1989; Talberth, Cobb, and Slattery, 2006; GPIAtlantic, 2007; Osberg and Sharpe, 2010; Anielski, 2001; United Nations, 2015).

Two thirds of these efforts mention sustainability in their main objective, and half mention “well-being” in the same sentence. All seven include accounts of at least some environmental degradation. Not surprisingly, all seven accounting systems describe their architecture as “top-down,” although one, the Alberta GPI, was a blend which incorporated public consultations for some selection of variables (Sharpe and Smith, 2005, p. 20).

Several sources are cited as providing heritage to the methodology underlying the top-down approaches of these six measures (International Human Dimensions Programme on Global Environmental Change, 2012), in partic-

ular the legacy of the Index of Sustainable Economic Welfare (Cobb, 1989) and the Genuine Progress Indicator (Talberth, Cobb, and Slattery, 2006) which evolved from it. In all but one case, the resulting measure is an index denominated in units of currency which represents the consumer expenditure portion of the GDP along with certain additions and subtractions to reflect normative classifications, non-market work, and environmental depletion. The exception (GPIAtlantic, 2007) is composed of two parts; one is an augmented GDP, like the others, while the second part is a set of indicators which cannot be aggregated into money-denominated accounts. We discuss other such indicators below.

Augmented monetary accounts are usually directly comparable with GDP or, rather, changes in each can be sensibly compared. Because of the emphasis on environmental capital which is typically being degraded overall, GPIs often show a decline in net product or a smaller rise in net product even when traditionally-measured GDP is rising substantially. Augmented accounts also sometimes focus on explicitly removing components of the GDP such as the economic activity generated by poor health or by damage from crime. According to the famous critiques in Kennedy’s speech on the subject (Kennedy, 1968), these should not be counted positively towards GDP.<sup>8</sup> Thus, augmented accounts are particularly valuable in the context of a strong existing focus on GDP growth and a widespread implicit assumption that other desirable outcomes are likely to be correlated with GDP. In this context, income-denominated augmented accounts wield rhetorical power in highlighting the differences resulting from more inclusive or more discerning coverage in the accounts.

Augmented monetary accounts also share the advantages with GDP of being a single (scalar) value and having a seemingly natural interpretation, and are thus easily digestible by both policy makers and the public. Inclusive accounting systems have therefore had some success in raising awareness that GDP growth has been an optimistic metric of progress or, more pointedly, in suggesting that existing policies are tailored more towards ensuring growth in GDP than furthering a more meaningful measure of progress or well-being.

To some extent, however, the criticisms relevant to the GDP apply also to GPIs and other augmented accounting systems, and they have received substantial critiques (e.g. Neumayer, 1999). Firstly, while redressing the exclusion of certain kinds of capital stocks and production, augmented ac-

---

<sup>8</sup>This idea, of course, fallaciously assumes that the GDP is *meant* to measure desirable things, while it can instead be considered an amoral accounting figure. We consider the meaning of augmented GDPs to be similarly obscure.



counts cannot include everything. In practice, they comprise only the extra components which the producers of the accounts have found a way to value in monetary terms. In fact, this is similar to the GDP in that the scope of the GDP has grown to incorporate new forms of production and capital when it becomes feasible to quantify their value (Bridgman et al., 2012; Coyle, 2015) but the standards of “feasibility” for national statistical agencies and international standards bodies are more stringent. Thus, the changes in the content of the GDP happen slowly. By contrast, the large but inevitably still incomplete addition of extra production values and losses comprising a potential augmented accounting system appears to the establishment as an especially arbitrary set. In highlighting the incompleteness of GDP, augmented versions of GDP can end up undermining their own legitimacy.

A second challenge faced by efforts to “correct” national income accounts is that revisions and augmentations are still based on market prices. Because directly observable prices are generally not available for the environmental and social goods which augmented measures include, shadow prices are imputed or extrapolated from observable markets. The link between accounting system and experienced human well-being is therefore not, qualitatively, improved by GPIs. An economic argument relating market or shadow prices to the scope of available choices — or to the subset made in the context of market transactions — and thus to human well-being, is tenuous, especially in light of all that is now known about the importance of social factors and non-market behaviour in explaining subjective well-being. We revisit these issues in [section 2.4](#). Regardless of how quantitatively important these methodological omissions and corrections may actually be, the approach of augmenting national income accounts based on “missing” or “inappropriate” components shares these possibly objectionable assumptions and methods with the GDP it is intended to supplant.

Augmented accounts which incorporate changes in stocks of capital usually omitted from the GDP face a particular challenge in assigning economic valuations to the stock. For instance, in principle it is important to include human capital changes, including what is gained in on-the-job learning and what is lost through unemployment, in a sum of national income. Similarly, changes to natural capital of many kinds, including renewable and non-renewable resources and broader environmental goods like the ecosystem service of a stable and hospitable climate, are necessary components in principle of any complete and meaningful set of accounts. However, evaluating a stock in commensurable terms with the value of a current market consumption good is problematic, above all, because of the intertemporal comparison of consumptions. Evaluations depend sensitively on a number

of very speculative predictions about the future, in addition to strong assumptions about how valuations vary with different mixes of consumption goods. These problems can be so deep and of such magnitude (See the Technical appendix of Stern, 2006) that the enterprise of evaluation is anyway dominated by *ad hoc* or normative assumptions.

## 2.2 Collections of indicators

In this second class of indicators, numerous measures deemed important are left as separate indicators, not aggregated into a scalar value. Most systems for measuring progress and well-being do not consist of indicators which are naturally commensurable, i.e. which can naturally be converted into monetary equivalents and thereafter aggregated. As a result, many systems of indicators stop short of aggregation. In our survey, 38 initiatives fall into this category, which we denote “collections (or sets) of indicators” (see footnote on page 13). As mentioned earlier, individual measurable quantities in such a collection may be organized by topic or theme, possibly nested in more than one level, but no effort is made to aggregate the constituent values quantitatively into a summary index. In some cases (e.g., The Boston Foundation, 2014) there may also be stated targets associated with the level or with changes in the level of individual constituent values. Such goals, along with the practice of comparing observed values or trends to those of other jurisdictions, are ways of providing context to what may otherwise seem to be obscure statistical quantities to an average citizen or decision maker. Even when trends are calculated, “average trends” across indicators, for instance in the form of average percentage change or average percentage improvement, are not.<sup>9</sup>

The collections of indicator in our database come half from government agencies and half from non-government organizations. Nearly half (20 of 38) of them include the words “quality of life” in the phrase describing the purpose of the effort, and one third (ten) mention “well-being”. We have already examined the overall trend towards the latter terminology.

In general, we found little evidence of strong public or policy impact associated with most of our measures in this category, even though some of the private efforts have been recognized by governments as significant.

Some exceptions include the recent U.K. National Well-being Measures, which are being considered by a number of agencies (Office for National

---

<sup>9</sup>While one percentage may appear commensurable with every other percentage because they are both unitless, there is no meaningful sense in which a fractional increase in literacy should be added to a fractional increase in waste diversion from landfills, for instance.

Statistics, 2013), as well as some prominent regional indicators (Institute for Sustainable Development, United Way of Winnipeg, 2013; Federation of Canadian Municipalities, 2012). Interestingly, although it is now defunct, the Oregon Benchmarks system (Young, 2005a) had considerable uptake by policy makers, including for budget priorities (Young, 2005b, p.1) and accountability for programs and policies (Sharpe and Smith, 2005, p.27), in addition to the usage by non-government organizations. Also notable is the Vital Signs model, which has been reproduced in a number of countries and jurisdictions (Community Foundations of Canada, 2013).

The pros and cons of the unaggregated indicators approach follow largely from contrasting it with that of the augmented monetary accounts. First of all, without the need to be able to calculate commensurable values, there is no restriction on which indicators can be included. Therefore, any measurable quantity which has an intuitive or demonstrated relationship to a theoretical or popular conception of human well-being (or which represents a possible indicator of environmental assets) or sustainability, can be included in a compendium of indicators. Indeed, while bottom-up (consultative) or top-down (theoretical) processes and criteria can be used to guide the selection, it is ultimately the goals of succinctness and simplicity which likely constrain the size of efforts in this class.

One other criterion for making indicators easy to read is that constituent measures be defined so that higher values are unambiguously better or worse than lower values. This is not strictly necessary in the case of collections of indicators, so they also allow for a broader choice of relevant measures. For instance, the governmental Indicators of Well-being in Canada (Employment and Social Development Canada, 2014) include indicators on the geographic distribution of population, but there is no normative direction associated with such measures. Again, the articulation of goals or desirable trends as part of such a “dashboard” of indicators is a useful feature for both interpretation and for making the indicators a tool for policy accountability.

An obvious failing of un-aggregated dashboards is that they do not provide a headline indicator or summary trend, making them less accessible and less appealing for nonspecialist audiences. While this avoids the dumbing-down of the complex and multidimensional concept of progress, it may relegate such efforts to the role of data clearinghouses, rather than significant contributions to reframing and redirecting public conceptions and dialogue, or to providing accountability for policy, even though these latter goals may be those which motivate the effort. Exhaustive lists of indicators, or even of categories of indicators, can seem complex and overwhelming, and lose the appeal that is associated with a tangible or intuitive goal. Nevertheless,

there are plenty of efforts in our database with longevity of more than a decade.

Because statistical agencies already have large collections of data series describing many important aspects of economic, social, and environmental functioning, the extra step of choosing a subset of measures and building some thematic structure around them is sometimes not a great leap. As a result, a number of governments, as well as other groups, are undertaking such projects, with various branding around the themes described earlier.

An associated problem with collections of indicators, whether large or small, is that they give no indication about the relative importance of different constituent measures. This may be better than explicitly giving them all equal weight, but it does not provide a tool, even in principle, for a decision maker to make tradeoffs. Benefit-cost analysis is, ultimately, possible only in the case when all relevant considerations have been made commensurable at which point they can also be aggregated into a single objective function, as in the case of money-denominated accounts.

### 2.3 Indices

Due to the large number of moving parts in a collection of indicators, it could be said that such systems allow only for qualitative accounts of overall performance or progress. For instance, in order to assess how the situation has changed in Boston, one must appeal separately to a large number of measures and then describe patterns and generalizations among the many trends. An obvious and appealing approach is to create a summary index of these many measures, even if they are not naturally commensurable. This can be done by rescaling each indicator to a standard scale before taking a mean or geometric mean (e.g., as in the HDI, Human Development Report Office, 2013) across indicators. Other possible methods of aggregation include imposing thresholds for each indicator before aggregating binary values (e.g. Alkire and Foster, 2011) or, in order to compare multiple jurisdictions, one can, for each constituent indicator, rank the jurisdictions in order of desirable outcomes and then for each jurisdiction aggregate the rankings across constituent indicators.

The motivations for this kind of construction have already been stated earlier. Presenting a single number to summarize desirable outcomes provides an accessible, simple, and unambiguous report for diverse audiences. Condensing data into a scalar objective also allows in principle for benefit-cost calculations to be made by considering the likely effect of different hypothetical policy shifts on the index, although the usefulness of this is

of course limited by the meaningfulness of the index. The availability of a headline indicator can also serve as an organizing concept and accessible entry point into the details underlying it. Having a summary statistic available may increase the public value and prominence of an otherwise-overwhelming array of indicators (e.g. OECD, 2016b).

On the other hand, the drawbacks to building poorly-founded aggregates are obvious and cannot be overstated. While the problems involved in aggregating values even when they are commensurable can be extensive, there is little to be said in defence of the method of arbitrarily weighting and summing numbers taken from unrelated quantities. Most of the indices in our database use equal weighting between many constituent components, often scaled to their full range, potential full range, or interquartile range (BCStats, 2014) before aggregating.

Consider the calculation of a change in an index of progress. In principle, if the weighting is not accountable to some theory or empirics, then the change in the index could be made to take any value in the entire range of its constituent components, simply by choosing weights appropriately (Mizobuchi, 2014). Using a fixed set of *ad hoc* weights<sup>10</sup> is no more meaningful than any other contrived set, and the same is true, in principle, of the value of the resultant index. More importantly, the weights in an index, whether or not they make any ethical sense and whether or not they are, superficially, “equal,” carry normative weight. If the index is intended to represent a measure of social progress or sustainable development or something else desirable, those weights represent statements about the relative importance of different intermediate outcomes in the achievement of the overall objective; they impose a framework for the concept of well-being or progress. If the weights are not chosen meaningfully, what can be said of the index?

In theory, this problem with indices should make them unlikely to hold sway when subject to scrutiny. On the long term, any such system does undergo public scrutiny, especially if it becomes well-known, and it must be able to sustain such examination if policy makers are ever to advocate for use of the measure, or to be held accountable to its numbers. The Stiglitz report’s 9th recommendation is that statistical offices should provide information to empower others to aggregate across dimensions of life quality in various ways, to create a variety of indices (Stiglitz, Sen, and Fitoussi, 2009). In our database, 52% of indicators which fit purely into the “index” category have become defunct, as compared with 39% of the collections of indicators

---

<sup>10</sup>Usually the weights are superficially all equal, but the effective weighting depends on the units used or rescaling choice for each component before aggregation.

which have chosen not to aggregate their components (see [Figure 4](#)).

The Canadian Index of Wellbeing (CIW, 2014) might serve as an example of these challenges. It is a collection of 64 measures organized into eight domains, one of which is environment related. The flagship product, however, is an index which equally weights the components to produce a scalar summary value. CIW went to great efforts to promote the index, during and after twelve years of development, but its impact has been modest and some original backers are no longer involved. Despite a careful effort to select and validate the constituent measures, the CIW is an index which has no natural units and is clearly the result of extensive judgments. These characteristics represent challenges for intuitively interpreting quantitative shifts in the index, and therefore obstacles to gaining support from those who are accountable for policy outcomes. Mizobuchi, (2014) discusses the problems with unaccountable index weights and reviews some other recent approaches to weighting.

Despite these problems, great resources have been spent on building such indices, and we classify 35 out of 82 measures in our database as indices. Some, such as the United Nation’s Human Development Index (Human Development Report Office, 2013), an index built from only four constituents, have considerable longevity and despite plenty of controversy, may be said to have had an effect on policy by providing some scalar alternative to measuring development with GDP.

Possibly the most interesting development in the measurement of all-encompassing well-being metrics is the rise of indices which aggregate their non-commensurable indicators using empirically-accountable weights. One instance is Bergheim, (2010)’s *Progress Index*, which uses estimates of the relationship between components of the HDI in a panel cointegration approach to derive aggregation weights. However, Bergheim eschews the use of subjective well-being as a guide to weights, thus missing an opportunity for them to represent normative contributions to well-being. By contrast, in nearly every other case of empirically-determined weights in our database, the weights are based on the importance of constituent indicators in models accounting for life satisfaction. We leave their discussion for later in the next section, in which we treat more generally the case of measures focused explicitly on subjective well-being.

## 2.4 Measures of subjective well-being

The fourth category in our taxonomy accounts for just 12 measures in our database; however, as is evident in [Figure 4](#), these subjective well-being in-

dicators comprise a growing fraction of progress metrics. Conceptually, the simplest of these is a single measure of cognitive evaluation of life, averaged at the country level, published since 2012 as part of the World Happiness Report (Helliwell, Layard, and Sachs, 2015). These data come directly from the “Cantril Ladder” question<sup>11</sup> in Gallup’s World Poll. Because the population-weighted country-average of the response to this question, appropriately translated for over 150 countries, is a single indicator, the issues of complexity and aggregation which plague collections of indicators, indices, and money-denominated accounts, including the GDP, are absent. This indicator of overall well-being attracts worldwide press each year.

Only slightly more complex, conceptually, are two simple indices with meaningful aggregation strategies. The earliest innovation of this sort is Veenhoven’s (1996) *Happy life expectancy*, which is a national-level measure computed simply as the product of life expectancy and average happiness measured on a unit scale. Veenhoven provides a review of the problems with the aggregation inherent in existing indices of quality of life at the time, and justifies a reliance on SWB (and life expectancy) as an apparent, or observed, outcome without the need to guess at the contributing factors to high quality of life. In a similar vein, the Happy Planet Index combines satisfaction with life (SWL<sup>12</sup>) with two other national-level measures in an intuitive way (New Economics Foundation, 2012).

Another four indicators in our SWB category are also indices but are focused exclusively on sets of subjective outcomes (New Economics Foundation, 2009; Gallup-Healthways, 2014; Centres for Disease Control and Prevention, 2014; Australian Unity, 2010). For instance, the Australian Unity Wellbeing Index 2010 is based on a series of survey questions asking about respondents’ satisfaction with different domains of life. These are averaged together to create overall individual and community well-being indices. However, the choice of equal weighting for this index raises the same conceptual problems as we have described for other indices.

---

<sup>11</sup>This is worded: “Please imagine a ladder with steps numbered zero at the bottom to ten at the top. Suppose we say that the top of the ladder represents the best possible life for you and the bottom of the ladder represents the worst possible. If the top step is 10 and the bottom step is 0, on which step of the ladder do you feel you personally stand at the present time?” Helliwell et al., (2010) found that averages of this question had a closely similar pattern of explanatory factors as the more standard life satisfaction question (given below, on page 33) recommended for national statistical agencies (Stone, Mackie, et al., 2014; OECD, 2013).

<sup>12</sup>SWL refers to the single specific question given on page 33, while the similar acronym SWB encompasses all subjective well-being measures, including the cognitive ones like SWL as well as those which assess affective states.



The three remaining measures in our SWB category are again also indices and were mentioned as an important innovation at the end of [section 2.3](#). These incorporate objective information but aggregate it into scalar indices using weights derived from empirical estimates. They can thus be seen as synthetic estimates of SWB, using relevant and available indicators thought to be important or shown to be important for experienced well-being. We dub this approach the *synthetic index of SWL*. The Legatum Prosperity Index (Lind, 2014) uses a model explaining variation in Gallup World Poll’s Cantril Ladder responses to determine weights in the first level of aggregation in its 8-domain index. Similarly, a regional analysis in France carried out by Bigot et al., (2012) aggregates 11 dimensions thought to represent aspects of well-being based on the coefficients in a regression for the answer to the question, “In your current life, do you feel happy: never, occasionally, quite often, very often?” Lastly, the Economist’s *Quality of Life Index* is aggregated using weights derived from a regression for the life satisfaction question (The Economist Intelligence Unit, 2005).

These SWB-weighted indices benefit from bearing a headline scalar (i.e., aggregated) index, yet because their weighting scheme is less a matter of arbitrary choice, they do not suffer as much from the drawbacks related to accountability and theoretical foundation. Meaningful weights quantitatively justify the inclusion of component indicators, thus fostering interest in those more specific indicators by people wishing for detail and policy insights. We discuss this approach further in the sections to follow.

## 2.5 Blurred boundaries

Naturally, some progress measures blur the boundaries of the categories we have set out, above. For instance, Maryland’s GPI incorporates measures such as the Gini, a common index of income or wealth inequality, into their dollar-denominated GPI (Maryland Department of Natural Resources, 2012). There is no sensible monetary value to the Gini, so while some components of their index are more legitimately converted to a money-denominated cost, others are incorporated in an entirely *ad hoc* way, in the interest of ending up with a dollar-denominated index. Such efforts cannot truly be called augmented systems of accounts, but they are still indices. Like other indices, they lack a transparent rationale for aggregation and therefore suffer from the drawbacks of both categories.

Another hybrid indicator devised a different compromise between systems of monetary accounts and systems of unaggregated indicators. The *Atlantic GPI* consisted of two components — one an augmented-GDP account



and the other a collection of standalone indicators (GPIAtlantic, 2007). Rather than coercing all components of progress into the form of a monetary aggregate, those which could not be sensibly incorporated into the accounting system were kept as part of a separate list. This appears to be a more rigorous approach than that of the Maryland GPI.

Other systems blend the categories of index and collection of measures. OECD’s Better Life Index is fundamentally a small collection of measures, because the OECD does not impose any particular set of fixed weights onto the components but, as already discussed, it is presented as an index and, indeed, has a default set of weights which are uniform (See Mizobuchi, 2014; Kasparian and Rolland, 2012, for more on the sensitivity to these weights). Moreover, like the majority of indices in our database, the OECD’s has two levels of aggregation. Even though the eleven domains may not be aggregated, each domain measure is in fact an index created using arbitrary scaling and aggregation.

Lastly, we classified the nine indices described in [section 2.4](#) under both categories; that is, most of the subjective well-being measures come in the form of a special case of indices.

### 3 Selection of Indicators

Regardless of whether or not individual data series are aggregated into some form of scalar index, the focus and breadth of the content depends on the selected set of constituent measures. We may distinguish two general strategies for choosing a set of measurements to include in a broadly framed assessment of progress or well-being. First, a *theoretical approach* means a top-down specification of what constitutes desirable metrics or domains based on some preconceived principle or theory. For example, a combination of a capabilities approach (Sen, 1997) and a lifecourse events framework is used to explain the selection of indicators making up Employment and Social Development Canada’s (2014) national indicator. Another approach, *bottom-up*, is to build up a consensus of what is important to a people through democratic consultation or through flexible empirical criteria. These two paradigms are described next.

#### 3.1 Top-down or theoretical

We classify most of the measures in our database as top-down or partially top-down in their formulation. Having some theoretical framework may

serve to bolster the accessibility or intuitive appeal of a measure of well-being or progress, in addition to guiding its implementation. For instance, the New Economics Foundation's (2012) *Happy Planet Index* embodies a simple definition of progress: the ecological efficiency of production of happy life-years. It brings together average life satisfaction, life expectancy, and another accounting aggregate, the per capita "ecological footprint," in order to produce a novel national-level measure with an intuitive meaning and the potential for some popular appeal.

A more typical example of a conceptual framework driving a collection of measures is the Employment and Social Development Canada's (ESDC, 2014) *Indicators of Well-being in Canada*. While the ESDC claims to have undertaken extensive consultation with experts, the selection and design of indicators does not reflect any consultative process with the general public.

In fact, organizations providing indices or collections of indicators designed in a top-down way frequently cite the number of years of consultations and/or the number of experts consulted as a way to express the rigour of the selection process. These efforts vary in the degree and transparency of theoretical structure guiding the selection process, which inevitably involves plenty of judgment. Sometimes, for instance, an orientation such as Sen's (Sen, 1997) capabilities approach may be cited (Employment and Social Development Canada, 2014) as the framework behind a selection of indicators, even though another set of principles may arguably have led to the same outcome.

Michalos et al., (2011, p.6) of the *Canadian Index of Wellbeing* claim their own method is "pragmatic," which they define as a blend of top-down and bottom-up. "We proceed patiently, transparently, and flexibly, testing any ideas presented both against the hard evidence yielded by empirical research and against the common sense of the Working Group and as broad a constituency beyond it as our resources allow." Clearly, while well-intended, it is unlikely to be reproducible in its detail, nor is it strictly accountable to either empirics or theory. Regarded from the outside, and coupled with the large number of freely chosen parameters in the form of arbitrary weights, this is clearly a challenge for the index's legitimacy.

Ultimately, a framework underlying the definition of progress, of an index, or of an overarching objective will need to appeal to politicians and the public because it is policy makers and government agencies who are to pursue it, and the public who judges their priorities and rationale. In this context, the "pragmatic" method of the CIW and numerous other similar efforts faces a challenge because it has neither an easy-to-grasp theoretical basis nor any prominent, broad, public consultative process which could

legitimize it for a generation or more.

### 3.2 Bottom-up: democratic or empirical

Alternatively, a more empirical, or atheoretical, source of authority may underlie the design and justification of an index or set of indicators. In one version of this approach, consultation occurs not primarily with experts but instead with the population in the form of direct or consultative democracy. For instance, in 2010 and 2011 the UK Office of National Statistics (ONS) undertook an ambitious program to construct a national consensus definition of *national well-being*, under a new objective to “Measure what matters.” Rather than limiting their focus to traditional measures of national accounts and economic and social performance, or to new measures based on someone’s theoretical framework, the ONS ran a series of events around the country, as well as extensive online debates and several modes for comment submissions, in order to solicit opinions from the population about what is important in life, how to measure national well-being, and how to use such a measure. These ideas were consolidated into a set of indicators comprising 10 domains and 38 individual measures (Office for National Statistics, 2012).<sup>13</sup>

Another somewhat less transparent but also empirical approach to building a well-being measure may be useful for building an index. This is exemplified by an opportunistic outcome of the OECD’s WWW site, which is primarily a communication tool but also serves as a research tool because it records user interaction. The site allows WWW site visitors from all over the world to choose and adjust the weights used to aggregate a set of 11 indicators which are calculated for all OECD countries (Boarini and d’Ercole, 2013). If the sets of weights users choose for viewing are assumed to reflect their preferences for constructing the best possible measure of national well-being, then indices of overall well-being could subsequently be constructed

---

<sup>13</sup>The ONS ended up with several measures of subjective well-being from this process, including the cognitive life satisfaction question, as well as emotional states of happiness and anxiety. Because O’Donnell and Oswald, (2015) propose to privilege the SWB measures in an overall index of well-being for the use of policy makers, they suggest deriving weights for the various SWB questions using another national survey, in which respondents suggest how important each emotion is. These would result in democratically-chosen weights. However, in describing life satisfaction as an “emotion” and putting multiple SWB measures on the same footing, these authors ignore the physiological distinction between a cognitive (cerebral), all-encompassing evaluative task and the task of querying (mid-brain) emotional states. See OECD, (2013) for why life satisfaction should be privileged, and Kahneman and Krueger, (2006) for a contrary view, in which it should be ignored in favour of emotions.

using somewhat democratically-sourced average weights.

Both of the methods described so far rely on individuals' beliefs or intuitions about what is important for defining overall national well-being or progress. Interpreted as democratic decision making processes (less so in the case of the OECD site, where the process is hidden), these methods make sense for building a collective social objective or policy assessment framework.

However, in terms of building valid measures or informing desirable outcomes, it is not obvious that a democratic method would be better than "top-down" expert consultation. Moreover, if instead the question "What matters most to you?" has in fact the meaning, "Which factors do you believe contribute most to your overall well-being?" then the question is a matter of prediction rather than selection. That is, a problem with asking people which things matter most to them is that, having only lived one life, they may be mistaken about the causality behind the life quality which they currently experience. If the overall life quality is measurable with cognitive evaluations of life, then the question of which factors matter most is an empirical question which can be answered statistically, without relying on individuals' instincts, yet without a paternalistic imposition of values.

Thus arises the third "bottom-up," empirical way to construct a well-being measure, and one which can be used to create an index. In this approach, which is applied in three indices described in Section 2.4 on page 24, any candidate measures of important contributors to a good life may be proposed, but their weighting in an index of well-being is determined by their explanatory power in statistical analyses of life satisfaction. For instance, when explaining individual variation in life satisfaction using standard multiple regression techniques, important factors that can also be calculated at the individual level and therefore incorporated into the analysis will "rise to the top" with a strong coefficient, i.e. weight. Importantly, these methods are intended to estimate average effects on SWL of *marginal* changes or differences in the explanatory factors, rather than the overall contribution to well-being outcomes. However, indices of well-being and progress are valued for their ability to track changes, so that an index of national well-being based on weights derived from SWL regressions should embody a properly-weighted measure of more proximate objectives which could raise average life satisfaction.

In practice, this approach may not be as transparent as it first seems,<sup>14</sup>

---

<sup>14</sup>Moreover, some indicator projects have quite pointedly rejected this approach. Michalos et al., (2011, pp. 17-18) claim that SWB does not capture well-being broadly enough,

since one can model life satisfaction at different scales, such as that of the individual, the community, or the country, using a variety of model specifications.<sup>15</sup> Moreover, because the various constituent indicators of interest will themselves be correlated, the coefficients in an equation estimating SWL will vary depending on which explanatory factors are included. Typically, models accounting for variation in SWL at the individual or aggregate level are linear, “reduced form” relationships. As a result, econometrically, the derived weights in such estimates do not directly represent causality. A further nuance is that each weight thus derived will be estimated along with its own confidence interval; these will produce a corresponding confidence interval in the computed index.

Nevertheless, an extensive literature focused on understanding the determinants of SWL has produced a high degree of consistency in the patterns of coefficients on factors explaining levels and changes of SWL at the individual level, and of levels of SWL at the regional and national levels. A smaller and maturing literature is able to account for changes over time at the aggregated geographic scales. While the details will vary, and should do so, according to the population being described and to the set of factors used in a synthetic index of SWL, they are nevertheless empirically accountable and will be a matter of debate only according to universal statistical principles. In summary, the weights used in a synthetic index of SWL are neither intuitively obvious nor will they be fixed, but can be expected to evolve over time as the depth of data and the science improve.

As an illustration of how such weights might be generated, we use once again the example of national-level data comprising the OECD “Better Life Index.” While the 11 domains are actually each a sub-index constructed with arbitrarily scaled and equally weighted components, we can use an SWL regression in order to derive weights for aggregating those 11 domains for

---

and that using it for validation is “not feasible” for the CIW. These authors include “the state of the natural environment, ... consumption and production patterns, the earth’s carrying capacity, ... other people’s wars...” in their conception of “well-being”. We argue below that building such broad collections of indicators confuses two incommensurable objectives, and inevitably ends up in unaccountable and opaque metrics. Other prominent critiques of SWB include Sen’s (1999) objection that a grumbling rich man may report being happier than a contented peasant. However, newer evidence from global SWB data mitigates some of these concerns. Even crudely, it can be noticed that cognitive evaluations of life from over 150 countries, now published nearly annually in the World Happiness Reports, show enormous variation which is not consistent with fears about an overwhelming role for aspirations and adaptation. The newer understanding is reflected by the more recent stance of Sen and his coauthors (Stiglitz, Sen, and Fitoussi, 2009).

<sup>15</sup>In [section 4.1](#), we compare this approach with using SWL itself as a sole indicator.

the 34 OECD member countries. We use ordinary least squares to estimate the value of one domain, life satisfaction, using the others:

$$\text{SWL}_i = a + b X_i + \varepsilon_i \tag{1}$$

where the  $X_i$  represents a vector of the values of the other domains for country  $i$ ,  $a$  is a constant, and  $\varepsilon_i$  captures the country modeling error. The domain variables in  $X$ , along with the estimated coefficients and standard errors, are shown in [Table 1](#). For expository purposes, we show the standardized coefficient values which can be compared across dependent variables.

Column (1) of [Table 1](#) shows an estimate incorporating all domains, while column (2) restricts the equation to the domains with the largest coefficients. Not surprisingly to those familiar with typical findings in the life satisfaction literature, OECD’s measures for *Jobs*, *Community*, and *Health* all attract large coefficients in the estimate, although in this small sample at the national level, few coefficients are statistically significant.

While these results may not accord immediately with one’s intuition, they illustrate a number of benefits of empirically-accountable weighting. First, using a human-centred outcome to proxy for overall well-being can provide grounds for questioning prior assumptions about the most important contributors to progress and well-being. For instance, a common finding in the SWL literature is that the value of having meaningful employment is large compared with typical variations in household income; this is reflected in our toy estimate in [Table 1](#) and suggests the two measures might not be equally weighted in a conception of progress.

More surprising is the unexpected sign for the coefficient of the *Safety* index, which is composed of the (inverted) homicide and assault rates. A second benefit of building a synthetic index using SWL is that issues with the data may be brought to light. Assault and homicide are unlikely to be positively correlated with SWL if measured properly. More likely, there are measurement problems with these values as national indicators, in part because they depend heavily on the quality of reporting of crimes and may mask large heterogeneities within the population.<sup>16</sup> In any case, given our findings, it is unlikely that the *Safety* index is contributing in the intended way to any weighted version of the OECD index. Using empirically-accountable weights offers a chance to refine, or reject, the constituent measures themselves.

More generally, one of the major benefits of having access to a subjective

---

<sup>16</sup>The assault data come from a survey question on the Gallup World Poll. Risks and perceived impacts may vary considerably with social status, as well as type of assault, as OECD itself recognizes on the Better Life Index site.

	LifeSatisfaction	
	(1)	(2)
Jobs	<b>.88<sup>†</sup></b>	<b>.77<sup>†</sup></b>
	(.24)	(.13)
Community	<b>.18<sup>+</sup></b>	<b>.17<sup>+</sup></b>
	(.10)	(.088)
Health	.24	.16
	(.17)	(.14)
logSafety	<b>−.31<sup>*</sup></b>	<b>−.32<sup>†</sup></b>
	(.12)	(.094)
Housing	−.12	
	(.19)	
Income	−.095	
	(.12)	
Education	−.0003	
	(.15)	
Environment	−.013	
	(.12)	
Civicengagement	.007	
	(.11)	
WorkLifeBalance	−.058	
	(.091)	
obs.	34	34
$R^2(\text{adj})$	.746	.776

Significance: **0.1%<sup>†</sup>** **1%<sup>\*</sup>** **5%** **10%<sup>+</sup>**

**Table 1: Weights for OECD domains, inferred from a model of life satisfaction.** Observations are for the 34 OECD countries. Data are from the OECD, (2016b).

measure of overarching well-being is to help to ensure that objective measures of well-being dimensions are capturing what is intended, and that the best ones are used, based on their statistical relationship to subjective well-being. This is especially important because traditionally under-measured supports of well-being such as social connections, political voice, and security are still lacking robust and reliable objective measures; life satisfaction provides a guide in this search (*Recommendation 6* in Stiglitz, Sen, and Fitoussi, 2009).

## 4 Roles for subjective well-being and sustainability

We now proceed to discuss special roles for two possible components of overall indicators of well-being or progress: the subjective measures we have already emphasized, and indicators of environmental impacts and sustainability. These deserve special attention because of their prominence and simultaneously their relative novelty, and because more careful treatment of these domains may continue to shape the future of influential indicators of overall human progress.

The increased perception of these concepts as important public goals is reflected, for instance, in the Thailand national government’s “Green and Happiness Index” (Barameechai, 2007). It incorporated six domains, each with two to four sub-components, along the lines of other measures in the style of Bhutan’s Gross National Happiness. Only one of the six domains, “Surroundings and Ecological System”, was ostensibly environment related, and was in fact mostly focused on human needs. None of the domains assessed happiness *per se*. Rather, the term “happiness” refers to a collection of values assumed by its creators to be related to a good life, in a moral as well as psychological sense. However, as we have already discussed, new indicators are increasingly incorporating true measures of SWB as well as metrics of ecological health.

### 4.1 Subjective reports of well-being

The prominence and potential for SWB to contribute to human-centred metrics of progress has already been discussed, above, in the context of indices built with subjective indicators or SWB-derived weights. This growth is based on the idea that individuals can aggregate their experience in accordance with their own priorities and values in a way that no one else can and,



according to many proponents, indeed in a way to which any other concept of well-being should ultimately be accountable.

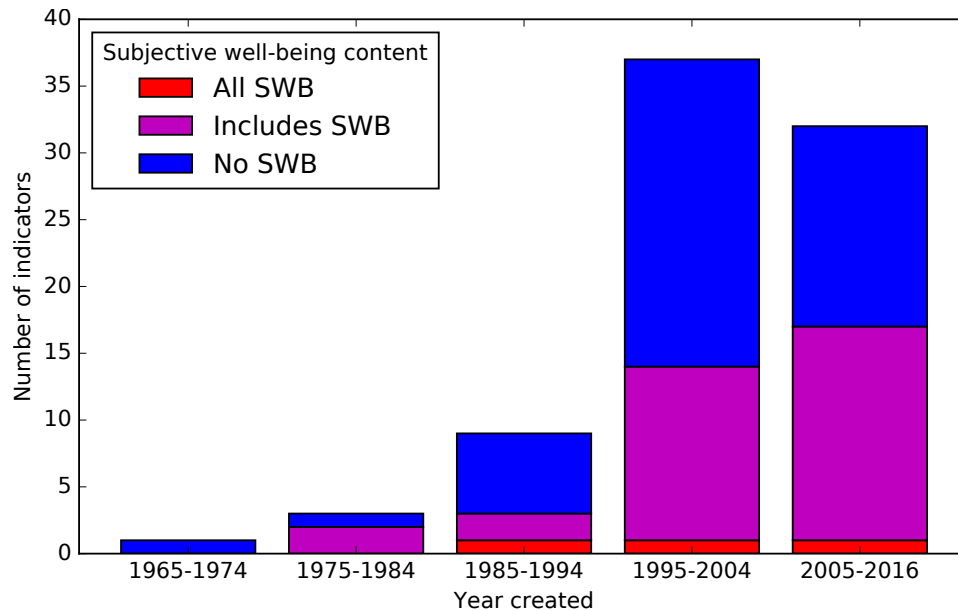
The idea that individuals have access, through reflection or feeling, to a meaningful and valid measure of overall well-being has prominent roots in the works of Aristotle and Bentham, to give examples from Western thought, but the understanding of these measures has advanced rapidly in recent years due to the increasing abundance of empirical data. As pointed out by Stiglitz, Sen, and Fitoussi, (2009), it is difficult to compare income over time in the face of technological change. It is also a great challenge to value publicly-provided individual services, as well as numerous other experiences which are not a result of choices. In addition, relying on choices as the last word on preferences, or welfare, is increasingly recognized as problematic. Cognitive evaluations of life, by contrast, in principle accommodate all these experiences and changes with the appropriate psychological weights.

Moreover, efforts by the U.S. National Academies (Stone, Mackie, et al., 2014) and in particular the OECD (OECD, 2013) have led to a standardization of SWB measurement. A single question forms the primary ingredient of these recommendations: *Overall, how satisfied are you with life as a whole these days, on a scale from 0 to 10, where zero means you feel “not at all satisfied” and 10 means you feel “completely satisfied”?* Plenty of evidence over the years shows that this measure exhibits stability within individuals, sensitivity to life conditions and changes, intuitive variation with material and other circumstances across the entire range of global national development levels, and even international and intercultural comparability (Helliwell et al., (2010) and Exton, Smith, and Vandendriessche, (2015), but see Lau, Cummins, and Mcpherson, (2005)). This comparability is one of the more remarkable properties, given decades of previous concern that subjective responses might reflect respondents’ limited aspirations and therefore represent, above all, their adaptation to their current circumstances, rather than an assessment of absolutely calibrated experience.<sup>17</sup>

Top-level policy makers, as well as numerous local and civic initiatives, are increasingly considering SWL as a basis for policy assessment. Evidence of this trend includes Prime Minister Cameron’s initiative in the U.K.

---

<sup>17</sup> Cummins and coauthors have used culture to explain some differences in life satisfaction responses, and this important question merits continued investigation. Various other influences on SWB have been identified and appear not to pose a problem for typical uses of the data. For instance, cultural biases in average response would not (necessarily) affect estimated coefficients in models of SWL within a population. Below, I propose the use of “synthetic” SWL measures, which are like an objective projection of SWL, and therefore especially robust for comparison across countries and cultures.



**Figure 5: Incorporation of subjective responses into measures of progress.**

(Cameron, 2010; UK Office of National Statistics, 2011; Dolan, Layard, and Metcalfe, 2011); Stiglitz, Sen, and Fitoussi’s report commissioned by President Sarkozy of France (Stiglitz, Sen, and Fitoussi, 2009); the OECD’s Better Life initiative OECD (OECD, 2015), OECD (OECD, 2016a); the World Happiness Reports (Helliwell, Layard, and Sachs, 2012, 2013, 2015); U.S. Federal Reserve chair Bernanke’s speech on well-being (Bernanke, 2010), and a growing number of more local initiatives in which subjective well-being is measured and targeted at the civic level.

However, the breadth of approaches evident in the range of extant measures of well-being and progress shows that there is still a lack of consensus on the appropriate role of SWB in measures of overall well-being. Should SWB contribute to progress measures as one dimension or category in an indicator or index? Is it, in this sense, comparable to objective measures of health? Does it make sense to aggregate the measure across individuals and is the pursuit of higher SWB, at the extreme, the very definition of human progress?

This debate has been articulated and developed by others (Veenhoven, 1996; Hagerty et al., 2001; Noll, 2004; Stiglitz, Sen, and Fitoussi, 2009; Fleurbaey and Blanchet, 2013; Helliwell, Huang, and Wang, 2015). We

report instead on the trends and innovations over time. [Figure 5](#) presents the results of classifying our database of indicators according to the inclusion of SWB measures of any kind. It shows not only the rise of indicators exclusively or primarily focused on SWB, but also the inclusion of SWB as an important component of sets of indicators and in indices. It is a reasonable prediction that prominent measures of well-being and progress will continue to embrace subjectively-assessed outcomes as an important standard of experience.

One may note that the synthetic SWL index proposed and discussed extensively in [Section 3.2](#) may in practice be quite similar to using straightforward averages of SWL itself. In principle, one might well prefer to focus on the SWL indicator itself, without building a synthetic index around it. While Helliwell, Huang, and Wang, (2015) argue for this approach, in preference to any sort of index, the same authors use a small set of other indicators to explain SWL and changes in SWL around the world (Helliwell and Wang, 2013), precisely in the way one would use SWL to organize and selectively emphasize a broader series of measures in an SWL-led set of indicators. In either case, differences and changes in mean SWL would be interpreted for policy purposes in terms of the differences and changes in a broader set of indicators, according to how those indicators can explain the observed SWL.

In fact, whether the average SWL itself or the synthetic version (i.e., deriving weights from a model of SWL) is presented as the headline indicator may be only a secondary consideration. The key and common features are an intuitive, human-centred, scalar well-being measure which is bundled with a set of more detailed indicators to support and explain it. Ideally, these other indicators are as objective as possible, allowing for the subjectivity to come only in determining what is important. In that case, a synthetic SWL index has the agreeable property that it is entirely constituted of objective measures, yet fully accountable to the content of experienced well-being. The synthetic index also has the advantage that it may be more easily calculated for subpopulations or small regions with limited sampling of SWL, but better data availability of the other indicators. Both SWL means and a synthetic SWL index can be calculated with confidence ranges, an important feature for any indicator whenever differences or changes in it may be considered (Helliwell, Huang, and Wang, 2015).

## 4.2 Environmental sustainability

There is overlap between the reviews of sustainability indices and well-being indices. Indeed, the concepts of sustainability are not altogether distinct

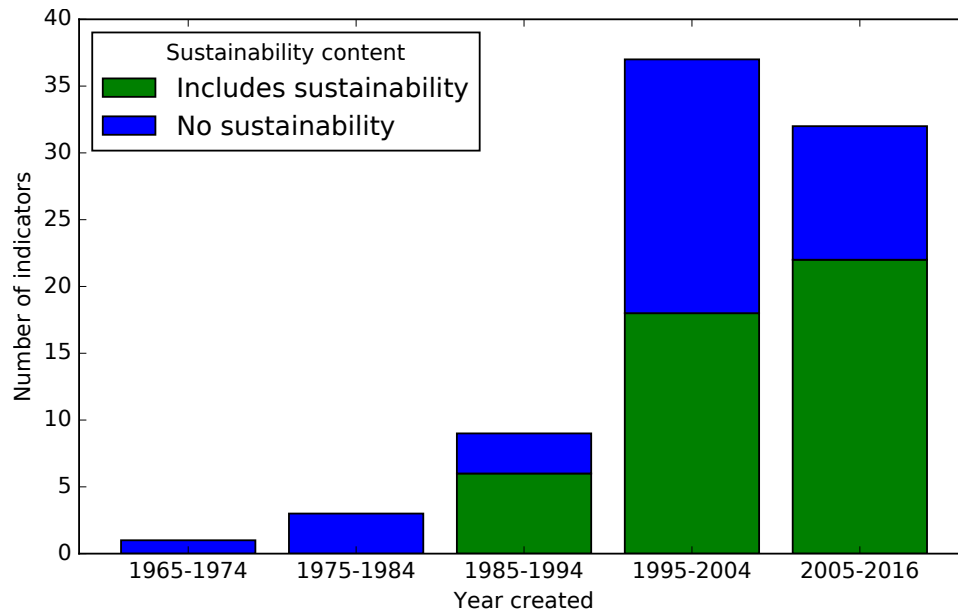
from well-being, in that definitions like that of the Brundtland et al., (1987) Report focus on human outcomes. Social, political, economic, and environmental dimensions of sustainability are often cited in the context of long-term human well-being. For this reason, sustainability and well-being are often conceived of and implemented together in measures of progress.

By contrast, the U.N.’s new Sustainable Development Goals contain elements which point more towards “strong sustainability” — that is, in which long-term human well-being is not the only objective. Behind this distinction is the additional issue of the extent to which current well-being experienced by humans incorporates people’s expectations and beliefs about future outcomes (Hall, Barrington-Leigh, and Helliwell, 2011). In an ideal measure of progress, would it be sufficient to evaluate the interests of the current generation, or should society advocate separately for the interests of those in the future? The answer to this question is clearly the latter for at least the following reasons: (I) Even future selves of currently living individuals appear to require separate representation by society; individuals do not fully or optimally prepare for the future due in part to what may be called, in general, “impatience.”<sup>18</sup> (II) In the context of the environment, it is clear that people are willing to live well, by any measure, in the present even at the cost of depleting resources for sustaining well-being in the future. (III) The evidence around subjective well-being gives no empirical reason to believe that it is any less present-biased than our behaviour. (IV) Other species may have the right to some representation as well; rather than debate the modern philosophy here, it suffices to point out that even Bentham, (1789) advocated, for the purposes of devising policy, aggregating well-being over all individuals of all species which could experience pain and pleasure.

These considerations suggest that measures of progress must, even in principle, contain more than one dimension if they are to account for the present and the future: an exclusive focus on the current experience of humans or on the current factors which affect it cannot be expected to capture the factors important for future well-being as well. Nevertheless, these two objectives are not necessarily commensurable. We explore whether devised indicators of progress have attempted to account for sustainability in addition to current human experience. [Figure 6](#) quantifies the growth of attention to measures of ecological sustainability in the indicators in our database. Our database excludes measures which are ostensibly focused solely on sustainability or ecological health; therefore the rise in the frac-

---

<sup>18</sup>Plenty of evidence for this exists, and gives rise to such policies as forced or incentivised savings.



**Figure 6: Incorporation of environmental indicators into measures of progress.**

tion of indicators which incorporate these themes reflects a change in what people consider to comprise progress or well-being. Most measures that do not include some indicators of environmental sustainability are focused on individual well-being, with names like *life situation*, *human development*, *social health*, *well-being index or accounts*, *quality of life*, *better life*, etc. Most measures that put sustainability at the core of their measurement, but nevertheless aim to address progress or well-being broadly, focus on broader notions of societal well-being, such as *wealth*, *national progress*, or *social indicators*.

Our finding of increased incidence of environmental indicators in broad measures of progress might largely reflect the growing visibility and importance of environmental limits to human well-being, rather than any conceptually-founded advances in integrating measures of environment and well-being. We consider these ongoing challenges below.

## 5 Discussion

While existing literature provides a number of principles for the design of valid indicators, our primary focus has been to survey the extant range of

indicators of social progress (see Sharpe and Smith, 2005, for a significant earlier survey) in order to survey points of consensus and dimensions of diversity. In these initiatives, constituent indicators for representing human and societal well-being have been chosen based on intuition and consensus, based on theory, or based on empirical evidence. They have been built by governments, by private organizations, and by academics. They span geographic scales from small communities to the whole world, and they are presented in a number of different formats in order to balance the demands of accessibility and meaningful specificity. Below we summarize our observations of these dimensions, along with our evaluation and predictions.

## Methodology

Despite the common motivation of improving our ability to record levels and changes in well-being and progress, we find still a diversity of methods being implemented to attain this end. Hagerty et al., (2001) and Michalos et al., (2011) emphasize a number of more general desirable qualities and “validity” criteria for indices of well-being, quality of life, and progress. Yet there remains in evidence an opposition of advocates of indices and advocates of unaggregated collections of measures. We have emphasized empirical accountability in our description of the relative merits of these. Implementations also differ in whether or not they define specific targets or goals in association with each indicator (as in Bhutan’s GNH or the Boston Indicators Project).

Broadly, our sample shows that efforts organized as indices have survived less well than those which remain as dashboards. This characterization is especially true of the money-denominated accounts, or augmented GDPs, which blossomed in the late 1990s and early 2000s, according to our sample, but have not survived well since then. The reasons for these differences could include the complexity and cost of aggregating, as opposed to merely collecting a set of periodically available statistics. The augmented GDP accounts are especially challenging to calculate, but their complexity is related to an opaqueness which is also likely to limit their usefulness.

In fact, since the turn of the century both the unaggregated indicators and the indices, even excluding the accounts, are becoming defunct as fast as new ones are being devised. In contrast, indicators we have classified as focused on subjective well-being are on the rise, and those which we found, even from the 1990s, are generally still in production.

One other trend we have pointed out is that indicators created by governments have tended to last longer than those from non-government orga-

nizations, and that the few initiatives from academia have struggled.

### Common themes

In our sample, there is considerable agreement with the general content of domains prevalent in many well-being and progress indicators. These are roughly consistent with the eight dimensions listed by Stiglitz, Sen, and Fitoussi, (2009, p. 14). Most of the following arise quite reliably in the indicators in our database: *material living standards; health; education; governance and civic participation; social connections, relationships, and community; environment; culture, accounts of time-use, and various forms of security* (e.g. Ura et al., 2012; The Canadian Index of Wellbeing, 2014b; Social Progress Imperative, 2014; Bigot et al., 2012; The Economist Intelligence Unit, 2005; Young, 2005b; World Health Organization, 2014).

Although some reviews have suggested that measures related to *political voice and governance* may be important mostly in politically unstable contexts (Hagerty et al., 2001), this is not borne out by the prevalence of indicators including these domains in stable and wealthy countries. It is also not consistent with the research using SWB, which suggests trust of and engagement with decision makers and institutions is universally valuable (e.g. as measured by perceived corruption in Helliwell et al., 2010).

**Education** Other differences of opinion may come down to measurement issues. *Education* is often assumed to have a fundamental role in defining quality of life measures, yet in studies accounting for the variance of SWB, education appears often to have primarily an instrumental role (e.g., The Economist Intelligence Unit, 2005, see also Table 1), especially when income and social capital are accounted for. Recently, there has been a successful series of initiatives in Canada, South Asia, and Latin America to shift the emphases and methods in education towards the promotion of non-cognitive skill in support of child well-being. It may be that holding education policy to account to well-being outcomes rather than test scores and labour markets could shift our education metrics to some more specific and relevant objectives.

**Inequality** Assumptions about the role of *inequality* for well-being have also been both softened and revived in the recent literature. Different kinds of inequality may matter to different outcomes, for instance for instrumental or indirect reasons. However, analyses using subjective well-being have often

found little role for income inequality in accounting for variation in well-being, once income itself is taken into account. It may be that the reasons to care about inequality are better measured more directly, for instance in the nature and quality of social capital and anonymous relationships, for appropriate levels of social insurance and investment, and so on.

Stiglitz, Sen, and Fitoussi, (2009) and others recommend that measures of material standard of living be recorded and communicated as they are experienced by households, i.e. rather than as aggregated (mean) income. Better representing the distribution in this way would more directly address some reasons for concern with inequality. More generally, a cross-cutting theme in recommendations for improved indicators is the shift from proxies to proximates, i.e. towards more direct measures of the outcomes that are intrinsically desirable.

**SWB** The summary literature preceding the present work (e.g., Veenhoven, 1996; Hagerty et al., 2001; Stiglitz, Sen, and Fitoussi, 2009; Fleurbaey and Blanchet, 2013) has emphasized subjective evaluation of life as an overall measure of well-being and as a statistical indicator of the relative importance of other domains and conditions in contributing to quality of life. It thus has a bearing on both content and methodology. Our focus on SWB has come from two directions: as an organizing concept and headline measure for human welfare or quality of life, and as a statistical tool to provide guidance on aggregation weights and even on what social and economic variables to incorporate in an indicator.

**Environment** By emphasizing that life satisfaction can be at the head of a set of indicators for well-being, we emphasize our conclusion (not new; see, e.g. Neumayer, 1999; Hall, Barrington-Leigh, and Helliwell, 2011) that it is overambitious to combine measures of current human well-being with metrics of environmental well-being. In addition, and for reasons on which we have elaborated here as well, there are deep problems turning environmental or sustainability assessments into a scalar index (Böhringer and Jochem, 2007). Thus, to address the broader concept of “progress,” as many in our database purport to do, requires more than one grouping of measures. In a theoretical sense, well-being might correspond to an objective function (i.e., that which is to be maximized), while maintenance of ecological standards acts as a set of constraints in devising optimal policy.



## Standardization

Eight years after the influential Sarkozy Commission, numerous well-being and progress measures have arisen around the world. Is it time for national statistical agencies or other organizations to begin to standardize these approaches?

Models like the GPI and the GNH have been copied and adapted a number of times. In addition, local communities implementing social indicators of well-being tend to look to others for successful models. This reproduction creates some *de facto* standardization. After some decades of increasing use in economics and statistics, the measurement of subjective well-being is now being standardized through studies and recommendations issued by OECD, (2013) and the U.S. National Academies (Stone, Mackie, et al., 2014). On the other hand, while the SDGs might be seen to represent a new global consensus on the meaning of progress, they are as yet only a vague guide to implementing indicators for it (Kroll, 2015), and earlier indices of development and poverty, even when produced by the U.N., have not been universally accepted.

There is enough in common among efforts for some degree of coalescence to a common framework. Moreover, there are considerable advantages to standardization. For one, a major motivation for many smaller jurisdictions to measure their progress is to compare outcomes with others. In addition, having an understandable structure which assures transparency is a necessity to be taken seriously by policy makers and their electorates, and the criteria for readability and transparency do not vary greatly from country to country. Another advantage to building a consensus on a broad set of progress measures is to reduce the redundancy within countries. For instance, within different departments of the Québec government, there are currently five different efforts to capture a general set of measures for well-being, progress, or sustainable development. Lastly, our understanding of human well-being and its determinants has evolved significantly, so a set of measures which properly captures the best science, while allowing for some continued evolution of understanding, is a feasible goal.

On the other hand, only a few countries have managed to put into place high profile progress measures and national government accountability in the spirit of the recommendations of Stiglitz, Sen, and Fitoussi, (2009): notably the U.K. and Bhutan, and shortly Germany.<sup>19</sup> At a smaller scale, there remains a large variety of methods characterizing the long-lasting and even

---

<sup>19</sup>Some other countries have related high-profile political initiatives with related branding, such as the United Arab Emirates, which appointed a Minister for Happiness.

influential measures in our database, meaning that by some criteria there may not be a clear winning strategy. Hall and Rickard, (2013) emphasize the benefits of the public deliberation itself, in particular in bottom-up processes, that comes from choosing or developing indicators, and Bergheim, (2015) provides detailed recommendations on this process. Above all, our knowledge of what *needs* to be measured to reflect experienced well-being appears to be far ahead of the *practice* of what is already actually being measured regularly and with high resolution. Except for the case of agencies which also have the funds to carry out their own data collection, only existing statistics are available for the construction of indices and indicators. As a result, the problem is not just conceptual but in most cases requires a significant political will and shift of resources to put new data collection into effect.

## 6 Conclusion

Our predictions are similar to our recommendations. One number will never encompass everything that needs to be monitored, but it can act as an organizing concept, provide some accountability to the selection of and emphasis given to other more specific indicators, and present a digestible face to the overall measurement objective. Unfortunately, existing indices with many components tend to be shorter lived (see Figure 4) and to lack transparency.

Overall, it appears that a collection of quantifiable conditions aggregated according to their contribution to SWL, when that contribution can be assessed, is a likely future trend for comprehensive, accountable, and easy to understand measures of well-being. Such a set would get at the core elements of indicators composed through other means, with some components such as education and economic inequality possibly playing a noticeably smaller role than they tend to now, but these are likely to be replaced by more proximate and detailed measures of social capital, civic participation, and meaningful work which appear to play direct roles in sustaining experienced well-being.

These human well-being measures will become generally more separated from systems of accounts which track environmental indicators, many of which will not be subject to aggregation in any meaningful way. Lastly, the environmental, social, human, and physical assets which support well-being and environmental integrity should and will continue to be tracked separately, and with increasing breadth and detail. Some of these assets may be commensurable in the way we have described for money-denominated

accounts, but such aggregation will never be complete.

## Recommendations

Based on (i) the fundamental index aggregation issues we have described, (ii) the need for empirical and conceptual accountability to social objectives, (iii) the benefits of being able to communicate a measure of progress with single, scalar value, and (iv) the distinct conceptual objectives of currently experienced human well-being versus long-term and non-human sustainability, we propose that the strongest model of indicators for progress appears to involve a separation into two components.

One component would be an SWL-weighted index, in which public consultations and available science could be used to suggest and select component indicators of social, economic, political, institutional, and other factors. A series of weights derived from the most appropriate models of SWL data would be available and the resulting index would serve as a headline indicator of well-being, possibly along with SWL itself.

Quite separately, an unaggregated system of indicators for environmental conditions and changes is necessary to account for the varied and non-commensurable threats to environmental sustainability. While this topic is not the focus of our study, it is clear that the full breadth of such indicators cannot be sensibly combined into a measure of well-being (Böhringer and Jochem, 2007). Instead, from a policy point of view, targets for its metrics should be seen more as constraints imposed on a society seeking human well-being.<sup>20</sup>

This dual system of indicators captures the components of progress and well-being, while leaving numerous issues of temporal integration and discounting up to policy. There may be some continued role for systems of accounts which attempt to monetize investments into social, human, and ecological capital. However, our survey has shown that these inevitably end up being contrived and incomplete. Moreover, they are likely to prolong the focus on GDP-like measures as the primary proxies for overall well-being, and thus forestall a progressive shift to broader and more human-centred objectives, such as those guided by observations of SWL. Ultimately, the

---

<sup>20</sup>This is unsatisfactory in light of a direct concern for non-human well-being, which is likely to become more explicit through law and policy over time. However, we would argue that these other forms of welfare should also stand apart, possibly next to human well-being, rather than being incorporated into broad sets of measures of the sustainability of environmental supports and systems. There is a range of philosophical approaches to such questions, but our criteria, above, are largely practical.

fact that GDP does not capture human well-being should not be confused with the fact that GDP does not account for environmental disinvestment.

Instead, when imputed monetary values can be assigned to environmental services, they can be made available alongside the relevant metrics for those services, individually. Adding up environmental deficits and subtracting them from the GDP tells us nothing about tradeoffs in overall investment unless the existing GDP is taken as a measure of valuable investment in the first place. For all the reasons motivating this present review, and indeed the social indicators movement from its outset, that is not an acceptable assumption.

## References

- Alkire, Sabina and James Foster (2011). “Counting and multidimensional poverty measurement”. *Journal of Public Economics* 95.7–8, pp. 476–487. ISSN: 0047-2727. DOI: <http://dx.doi.org/10.1016/j.jpubeco.2010.11.006>.
- Anielski, Mark (2001). *The Alberta GPI Blueprint: The Genuine Progress Indicator (GPI) Sustainable Well-Being Accounting System*. Tech. rep., pp. 1–126.
- Australian Bureau of Statistics (2014). *Measures of Australia Progress, 2013: What is MAP*. URL: <http://www.abs.gov.au/ausstats/abs@.nsf/Lookup/1370.0main+features672013> (visited on 04/09/2014).
- Australian Unity (2010). *Wellbeing*. Tech. rep.
- Baramuechai, Juthamas (2007). “The green and happiness index”. In: *International Conference on Happiness and Public Policy, Bangkok*, pp. 18–19.
- BCStats (2014). *BC Regional Socio-Economic Profiles & Indices*. URL: <http://www.bcstats.gov.bc.ca/StatisticsBySubject/SocialStatistics/SocioEconomicProfilesIndices.aspx> (visited on 03/29/2014).
- Bentham, Jeremy (1789). *An introduction to the principles of morals and legislation*. Reprinted 1948, Oxford, UK: Blackwell.
- Bergheim, Stefan (2010). “The Progress Index”.
- (2015). *Ten Steps to Success Quality of Life Processes – A Manual*. Tech. rep. Zentrum für gesellschaftlichen Fortschritt.
- Bernanke, B. S. (2010). *The Economics of Happiness*. URL: [www.federalreserve.gov/newsevents/speech/bernanke20100508a.htm](http://www.federalreserve.gov/newsevents/speech/bernanke20100508a.htm) (visited on 12/2011).
- Bigot, Régis et al. (2012). “L’évolution du bien-être en France depuis 30 ans”.
- Boarini, Romina and Marco Mira d’Ercole (2013). “Going beyond GDP: An OECD perspective”. *Fiscal Studies* 34.3, pp. 289–314.
- Böhringer, Christoph and Patrick E.P. Jochem (2007). “Measuring the immeasurable — A survey of sustainability indices”. *Ecological Economics* 63.1, pp. 1–8. ISSN: 0921-8009. DOI: <http://dx.doi.org/10.1016/j.ecolecon.2007.03.008>.
- Bridgman, Benjamin et al. (2012). “Accounting for household production in the national accounts, 1965–2010”. *Survey of Current Business* 92.5, pp. 23–36.
- Brundtland, Gro et al. (1987). “Our Common Future (“Brundtland report”)”.

- Cameron, Prime Minister David (2010). *A transcript of a speech given by the Prime Minister on wellbeing on 25 November*. URL: <http://www.number10.gov.uk/news/pm-speech-on-well-being> (visited on 12/2011).
- Central Statistics Office (2012). *Measuring Ireland Progress: 2011*. Tech. rep., pp. 1–102.
- Centre for Bhutan Studies (2015). *Gross National Happiness Index Explained in Detail*. URL: [http://www.grossnationalhappiness.com/docs/GNH/PDFs/Sabina\\_Alkire\\_method.pdf](http://www.grossnationalhappiness.com/docs/GNH/PDFs/Sabina_Alkire_method.pdf) (visited on 03/08/2015).
- Centres for Disease Control and Prevention (2014). *Health Related Quality of Life: About Us*. URL: <http://www.cdc.gov/hrqol/about.htm> (visited on 05/15/2015).
- Cobb, Clifford W (1989). “The index for sustainable economic welfare”. Ed. by HE Daly and JB (eds) Cobb, pp. 401–57.
- Community Foundations of Canada (2013). *Vital Signs*. URL: <http://www.vitalsignscanada.ca/en/about> (visited on 05/11/2014).
- Cooke, Martin (2005). *The First Nations Community Well-Being Index (CWB): A Conceptual Review*. Tech. rep. Indian Affairs and Northern Development, pp. 1–24.
- Coyle, Diane (2015). *GDP: A Brief but Affectionate History*. Princeton University Press. ISBN: 9781400873630.
- Department for Environment Food & Rural Affairs (2013). *Sustainable Development Indicators*. Tech. rep., pp. 1–99.
- Dolan, P, R Layard, and R Metcalfe (2011). “Measuring Subjective Well-being for Public Policy”. *Office for National Statistics paper*.
- Employment and Social Development Canada (2014). *Indicators of Well-being in Canada: About this site*. URL: [http://www4.hrsdc.gc.ca/c.4nt.2nt@-eng.jsp?cid=14%5C#header\\_1](http://www4.hrsdc.gc.ca/c.4nt.2nt@-eng.jsp?cid=14%5C#header_1) (visited on 05/11/2014).
- European Statistical System (2012). *Measuring Progress, Well-being and Sustainable Development: The response of the European Statistical System to the report of the Stiglitz-Sen-Fitoussi Commission and to the ‘GDP and beyond’ Communication of the European Commission*. Tech. rep., pp. 1–4. DOI: [10.2785/21299](https://doi.org/10.2785/21299).
- Exton, C., C. Smith, and D. Vandendriessche (2015). “Comparing Happiness across the World: Does Culture Matter?” *OECD Statistics Working Papers*. DOI: <http://dx.doi.org/10.1787/5jrqqppzd9bs2-en>.
- Federation of Canadian Municipalities (2012). *Quality of Life Reporting System: About the program*. URL: <http://www.fcm.ca/home/programs/quality-of-life-reporting-system/about-the-program.htm> (visited on 05/11/2014).

- Fleurbaey, M. and D. Blanchet (2013). *Beyond GDP: Measuring Welfare and Assessing Sustainability*. Oxford University Press. ISBN: 9780199344154.
- Gallup-Healthways (2014). *State of American Well-Being: 2013 State, Community and Congressional District Analysis*. Tech. rep., pp. 1–38.
- GPIAtlantic (2007). *Community GPI*. URL: <http://gpiatlantic.org/community.htm> (visited on 03/08/2015).
- Hagerty, Michael et al. (2001). “Quality of Life Indexes for National Policy: Review and Agenda for Research”. *Social Indicators Research* 55.1, pp. 1–96. ISSN: 0303-8300. DOI: [10.1023/A:1010811312332](https://doi.org/10.1023/A:1010811312332).
- Hall, Jon, Christopher Barrington-Leigh, and John Helliwell (2011). “Cutting through the Clutter: Searching for an Over-Arching Measure of Well-Being”. *DICE Report* 8.4, pp. 8–12.
- Hall, Jon and Louise Rickard (2013). *People, progress and participation: How initiatives measuring social progress yield benefits beyond better metrics*. Gütersloh: Bertelsmann Stiftung.
- Helliwell, J.F. et al. (2010). “International evidence on the social context of well-being”. In: *International Differences in Well-Being*. Ed. by E. Diener, J.F. Helliwell, and D. Kahneman. Oxford University Press, pp. 213–229.
- Helliwell, John, Haifang Huang, and Shun Wang (2015). “The geography of world happiness”. In: *World Happiness Report 2015*. Ed. by John Helliwell, Richard Layard, and Jeffrey Sachs. Sustainable Development Solutions Network. Chap. 2, pp. 12–41.
- Helliwell, John, Richard Layard, and Jeffrey Sachs (2012). *World Happiness Report*. URL: <http://worldhappiness.report/ed/2012/>.
- (2013). *World Happiness Report 2013*. URL: <http://worldhappiness.report/ed/2013/>.
- (2015). *World Happiness Report 2015*. URL: <http://worldhappiness.report/ed/2015/>.
- Helliwell, John and Shun Wang (2013). “World happiness: Trends, explanations, and distribution”. In: *World Happiness Report 2013*. Ed. by John Helliwell, Richard Layard, and Jeffrey Sachs. Sustainable Development Solutions Network. Chap. 2, pp. 58–37.
- Hicks, Peter (2012). “A framework for citizen-centred social statistics and analysis”. *Queen’s University School of Public Policy Working Paper* 48.
- Human Development Report Office (2013). *Human Development Report 2013: Technical Notes*. Tech. rep., pp. 1–8.
- Institute for Sustainable Development, United Way of Winnipeg (2013). *About Peg*. URL: <http://www.mypeg.ca/about> (visited on 05/10/2014).

- International Human Dimensions Programme on Global Environmental Change (2012). *Inclusive Wealth Report: Policy Recommendations*. URL: <http://www.ihdp.unu.edu/article/read/iwr-policy>.
- Jacksonville Community Council Inc (2014). *Jacksonville Quality of Life Indicators*. URL: <http://www.jcci.org/#!indicators/c1e39> (visited on 03/08/2015).
- Kahneman, D. and A.B. Krueger (2006). “Developments in the measurement of subjective well-being”. *Journal of Economic Perspectives* 20.1, pp. 3–24.
- Kasparian, Jérôme and Antoine Rolland (2012). “OECD’s “Better Life index”: can any country be well ranked?” *Social Applied Statistics*, pp. 1–8.
- Kennedy, Robert F. (1968). Speech on 18 March 1968 at the University of Kansas. URL: <http://www.jfklibrary.org/Research/Research-Aids/Ready-Reference/RFK-Speeches/Remarks-of-Robert-F-Kennedy-at-the-University-of-Kansas-March-18-1968.aspx> (visited on 03/09/2015).
- Kroll, Christian (2015). “Sustainable Development Goals: Are the rich countries ready?” *Spotlight Europe*.
- Land, Kenneth C. and Alex C. Michalos (2016). “Fifty years after the social indicators movement: has the promise been fulfilled? An Assessment and an Agenda for the Future”. *this issue*.
- Lau, Anna, Robert Cummins, and Wenda Mcpherson (2005). “An Investigation into the Cross-Cultural Equivalence of the Personal Wellbeing Index”. *Social Indicators Research* 72.3, pp. 403–430. ISSN: 03038300.
- Lind, Niels (2014). “Legatum Prosperity Index”. In: *Encyclopedia of Quality of Life and Well-Being Research*. Springer, pp. 3529–3530.
- Maryland Department of Natural Resources (2012). *Maryland Genuine Progress Indicator*. URL: <http://www.dnr.maryland.gov/mdgpi/> (visited on 03/09/2016).
- Michalos, Alex et al. (2011). *Technical Report 1.0*. Tech. rep. Waterloo, ON: The Canadian Index of Wellbeing and University of Waterloo.
- Ministry of Social Development (2010). *2010 The social report: Social Indicators for New Zealand*. Tech. rep.
- Mizobuchi, Hideyuki (2014). “Measuring world better life frontier: a composite indicator for OECD better life index”. *Social Indicators Research* 118.3, pp. 987–1007.
- Neumayer, Eric (1999). “The ISEW – not an Index of Sustainable Economic Welfare”. *Social Indicators Research* 48.1, pp. 77–101. ISSN: 0303-8300. DOI: [10.1023/A:1006914023227](https://doi.org/10.1023/A:1006914023227).



- New Economics Foundation (2009). *NAW-About*. URL: <http://www.nationalaccountsofwellbeing.org/> (visited on 06/03/2014).
- (2012). *The Happy Planet Index: 2012 report*. Tech. rep.
- Noll, Heinz-Herbert (2004). “Social Indicators and Quality of Life Research: Background, Achievements and Current Trends”. In: *Advances in Sociological Knowledge*. Ed. by Nikolai Genov. VS Verlag für Sozialwissenschaften, pp. 151–181. ISBN: 978-3-8100-4012-1. DOI: [10.1007/978-3-663-09215-5\\_7](https://doi.org/10.1007/978-3-663-09215-5_7).
- O’Donnell, Gus and Andrew J. Oswald (2015). “National well-being policy and a weighted approach to human feelings”. *Ecological Economics* 120, pp. 59–70. ISSN: 0921-8009. DOI: <http://dx.doi.org/10.1016/j.ecolecon.2015.09.021>.
- OECD (2016a). *Better Life Initiative: Measuring Well-being and Progress*. URL: <http://www.oecd.org/progress> (visited on 03/2016).
- (2011). *How’s Life? Measuring Well-being*. Tech. rep., pp. 1–286. DOI: [10.1787/9789264121164-en](https://doi.org/10.1787/9789264121164-en).
- (2013). *OECD Guidelines on Measuring Subjective Well-being*.
- (2015). *How’s Life?: Measuring well-being*. OECD Publishing.
- (2016b). URL: <http://oecdbetterlifeindex.org/>.
- OECD Better Life Initiative (2014). *Executive Summary*. Tech. rep.
- Office for National Statistics (2012). *Measuring National Well-being: Report on consultation responses on proposed domains and measures*. Tech. rep., pp. 1–40.
- (2013). *Wellbeing Policy and Analysis : An Update of Wellbeing Work across Whitehall*. Tech. rep., pp. 1–40.
- Osberg, Lars and Andrew Sharpe (2010). *Beyond GDP: Measuring Economic Well-Being in Canada and the Provinces, 1981-2010*. Tech. rep. Centre for the Study of Living Standards, pp. 1–85.
- Perry, Bryan (2009). *Non-income measures of material wellbeing and hardship: first results from the 2008 New Zealand Living Standards Survey, with international comparisons*. Tech. rep. New Zealand Ministry of Social Development, pp. 1–76.
- Porter, Michael E., Scott Stern, and Michael Green (2014). *Social Progress Index 2014*. Tech. rep.
- Redefining Progress (2014). *About Redefining Progress*. URL: [http://rprogress.org/about\\_us/about\\_us.htm](http://rprogress.org/about_us/about_us.htm) (visited on 07/15/2014).
- Sen, Amartya K. (1999). *Development as Freedom*. Oxford University Press.
- Sen, Amartya Kumar (1997). *Choice, welfare and measurement*. Harvard University Press.

- Shackman, Gene, Ya-Lin Liu, and Xun Wang (2005). “Measuring a quality of life using free and public domain data”. *Social Research Update* 47.
- Sharpe, Andrew and Jeremy Smith (2005). *Measuring the Impact of Research on Well-being: A Survey of Indicators of Well-being*. Tech. rep. Centre for the Study of Living Standards, pp. 1–119.
- Social Progress Imperative (2014). *Social Progress Index - Methodology*. URL: <http://www.socialprogressimperative.org/data/spi/methodology> (visited on 07/25/2014).
- Social Progress Index* (2014). Social Progress Imperative. URL: <http://www.socialprogressimperative.org/data/spi> (visited on 03/08/2014).
- Social Weather Stations (2015). *Philippines Social Weather Report*. URL: <http://www.sws.org.ph/>.
- Statistics Finland (2014). *Well-being Indicators*. URL: <http://www.findikaattori.fi/en/hyvinvointi> (visited on 07/18/2014).
- Stern, N.H. (2006). *Stern Review on the Economics of Climate Change*. U.K. Treasury.
- Stiglitz, Joseph E, Amartya Sen, and Jean-Paul Fitoussi (2009). “Report by the Commission on the measurement of economic performance and social progress”. *Paris: Commission on the Measurement of Economic Performance and Social Progress*.
- Stone, Arthur A, Christopher Mackie, et al. (2014). *Subjective well-being: Measuring happiness, suffering, and other dimensions of experience*. National Academies Press.
- Talberth, John, Clifford Cobb, and Noah Slattery (2006). *The Genuine Progress Indicator*. Tech. rep.
- The Boston Foundation (2014). *The Boston Indicators Project: About the Project*. URL: <http://www.bostonindicators.org/about> (visited on 05/11/2014).
- The Canadian Index of Wellbeing (2014a). *About the Canadian Index of Wellbeing*. URL: <https://uwaterloo.ca/canadian-index-wellbeing/about-canadian-index-wellbeing> (visited on 05/11/2014).
- (2014b). *Domains*. URL: <https://uwaterloo.ca/canadian-index-wellbeing/our-products/domains> (visited on 05/11/2014).
- The Economist Intelligence Unit (2005). “The Economist Intelligence Unit quality-of-life index”. *The Economist*.
- The Netherlands Institute for Social Research (2015). *Life situation Index*. URL: <http://www.scp.nl/english/leefsituatie/>.
- Treasury Board of Canada (2010). *Canada’s Performance: Annual Report to Parliament 2009-10*. Tech. rep. Government of Canada, pp. 1–105.

- UK Office of National Statistics (2011). *National Statistician's reflections on the National Debate on Measuring National Well-being*. EPRINT: [www.ons.gov.uk](http://www.ons.gov.uk). URL: <http://www.ons.gov.uk/ons/guide-method/user-guidance/well-being/wellbeing-knowledge-bank/understanding-wellbeing/measuring-what-matters--national-statistician-s-reflections-on-the-national-debate-on-measuring-national-well-being.pdf> (visited on 12/2011).
- UNEP, UNU-IHDP (2015). *Inclusive Wealth Report 2012*. URL: <https://www.ihdp.unu.edu/article/read/iwr> (visited on 03/08/2015).
- United Nations (2015). *Inclusive Wealth Report 2012*. URL: <https://www.ihdp.unu.edu/article/read/iwr> (visited on 03/08/2015).
- United Nations Development Programme (2015). *Human Development Index*. URL: <http://hdr.undp.org/en/content/human-development-index-hdi> (visited on 03/09/2015).
- Ura, Karma et al. (2012). *An Extensive Analysis of GNH Index*. Tech. rep. Centre for Bhutan Studies, Under the Patronage of His Majesty the King, pp. 1–148.
- Veenhoven, Ruut (1996). “Happy life-expectancy”. *Social Indicators Research* 39.1, pp. 1–58.
- World Health Organization (2014). *WHO Quality of Life-BREF*. URL: [http://www.who.int/substance\\_abuse/research\\_tools/whoqolbref/en/](http://www.who.int/substance_abuse/research_tools/whoqolbref/en/) (visited on 05/15/2014).
- Young, Richard (2005a). *An Overview: Oregon Shines II and Oregon Benchmarks*. Tech. rep. Institute for Public Service and Policy Research, University of South Carolina.
- (2005b). *An Overview: Oregon Shines II and Oregon Benchmarks*. Tech. rep. Institute for Public Service and Policy Research, University of South Carolina.