Received: 31 May 2021 | Revised: 2 August 2021 | Accepted: 16 August 2021

DOI: 10.1002/sd.2234

RESEARCH ARTICLE

Sustainable Development WILEY

Toward better measurement of sustainable development and wellbeing: A small number of SDG indicators reliably predict life satisfaction

Ida Kubiszewski[1] | Kenneth Mulder[2] | Diane Jarvis[3] | Robert Costanza[1]

[1]Crawford School of Public Policy, The Australian National University, Canberra, Australian Capital Territory, Australia

[2]Long Trail School, Dorset, Vermont, USA

[3]College of Business, Law & Governance, James Cook University, Townsville, Queensland, Australia

Correspondence
Ida Kubiszewski, Crawford School of Public Policy, The Australian National University, Canberra, ACT, Australia.
Email: ida.kub@gmail.com

Abstract

In 2015, all 193 member states of the United Nations (UN) adopted the Sustainable Development Goals (SDGs). These 17 goals, 169 targets, and 232 indicators (including over 650 indicators if all the subdivisions are included) are intended to guide and improve sustainable wellbeing and life satisfaction for everyone on earth. Challenges include the fact that many indicators are not measured or reliably tracked in many countries, the cost of tracking is unclear, and no explicit overarching goal exists. To highlight some of the problems with this approach, we model life satisfaction (LS) survey scores by country, as a proxy for overall wellbeing, as the dependent variable against the official 232 SDG indicators. Using a constrained linear regression approach (LASSO), we identify a model that includes only 8 of the 232 indicators and explains 84% of the variation in LS. These eight indicators are proxies for economic, social, and environmental variables. We also cluster countries according to these indicators and LS showing correlation within geographical and cultural regions. We discuss these results with regard to the meaning and measurement of sustainable development vs. sustainable wellbeing and its relationship with LS and the SDGs. We recommend how these results can be used to prioritize goals and measurement efforts to create more meaningful and useful measures of sustainable wellbeing.

KEYWORDS

indicators, life satisfaction, sustainable development, Sustainable Development Goals (SDGs), wellbeing

1 | INTRODUCTION

The 2030 Agenda for Sustainable Development and the Sustainable Development Goals (SDGs) were adopted by 193 countries as a "blueprint to achieve a better and more sustainable future for all" (United Nations, 2020). This is the first time all countries around the world achieved consensus around a set of global goals that apply to all countries. Unlike the Millennium Development Goals (MDGs), which were focused on developing countries, the SDGs set the same set of goals for every country in the world.

The 17 goals are supported by 169 targets and 232 indicators. They encompass social, economic, environmental, and institutional aspects. Measuring these indicators has been described as an "unprecedented statistical challenge" by the President of the UN

General Assembly because of the complexity of the targets, their global nature, and occasional lack of agreed statistical definition (Lebada, 2016; MacFeely, 2020).

An additional issue is that official country data, instead of international data, is used as the primary source for populating the SDG indicators (MacFeely, 2020). This means that due to a lack of in-country resources, especially in developing countries, many indicators will remain unpopulated as they fall far outside the scope of what is able to be collected by national statistical offices (Kapto, 2019). This problem is further exacerbated by many of the indicators themselves having subdivisions (e.g., unemployment rate is subdivided into female and male unemployment, others are subdivided based on age groups or rural versus urban). Including all these subdivisions would yield more than 650 indicators. The cost of adequately measuring the 232 indicators across all countries has been estimated to be approximately $45 billion over the 15 years, or $3 billion a year (Badiee et al., 2016). But this is dwarfed by UN estimates of the cost of actually achieving the SDGs by 2030, which vary between $5 and $7 trillion dollars a year (Vorisek & Yu, 2020).

Even if measured, these 232 independent indicators pose a non-trivial problem of to coalesce into a single indication of progress within each target and goal (Mair et al., 2017). Also, no hierarchy (Schmidt-Traub et al., 2017) or overarching goal exists to help prioritize or bring these 17 goals together into a single measure of progress for the world that adequately addresses the synergies and trade-offs among the goals and how these vary across countries (Allen et al., 2016, 2017; Costanza, Daly, et al., 2016; Daly, 1973; Le Blanc, 2015; Nilsson et al., 2016; Sachs et al., 2019).

Even if all the SDG indicators were to be easily measured by all countries in the world, they would not provide a good measure of sustainable wellbeing due to omission of some important measures, misspecification of others, and an overall unbalanced distribution of social, environmental, and economic factors (Aksoy & Bayram Arlı, 2020; Costanza, Daly, et al., 2016; Giannetti et al., 2020). For example, although goal 10 is “Reduce Inequality,” the indicators associated with this goal do not translate into capturing progress toward reducing inequality, particularly within countries (Winkler & Satterthwaite, 2017).

The overarching goal of the SDGs is, in theory “sustainable development.” But the term “development” is often interpreted as continuing the development trajectories of the “developed” countries based on GDP growth, rather than the more basic meaning of improvement in quality (Daly, 1996). We argue that global society's overall goal should be sustainable wellbeing, which depends on the wellbeing of our ecological life support system (Bai et al., 2016; Costanza et al., 2018; Helne & Hirvilammi, 2015; Kubiszewski et al., 2013). The term “sustainable wellbeing” is more consistent with the comprehensive nature of the SDGs than “sustainable development,” and recognizes that “development” is often misinterpreted to mean “growth.” Growth is one means to the end of maintaining and enhancing wellbeing (and strongly relates to SDG #8 “Decent work and economic growth”) but is not the end in itself.

Societal wellbeing is difficult to measure directly. Many wellbeing indicators and indices exist, using either objective or subjective variables, or a combination (Costanza et al., 2014). Subjective wellbeing, in the form of self-reported life satisfaction (LS), is one component of wellbeing. Some have argued that improving LS should be the primary goal of social policy, since LS integrates across a range of conditions that affect people's lives (Dolan et al., 2011; Layard, 2006).

We recognize that LS is only one component of overall sustainable wellbeing. There are well known issues and inconsistencies with using LS as a proxy for overall wellbeing, including: cultural differences (Graham & Markowitz, 2011), varying perceptions of reality (Ambrey et al., 2014; Kubiszewski et al., 2018), values held by communities (Kubiszewski, Jarvis, et al., 2019), and personality differences (Kubiszewski et al., 2020; Soto, 2015). Also, individuals do not necessarily have access to all the information about what impacts their wellbeing, especially aspects that are beyond their direct perception (i.e., regulating ecosystem services) or far in the future (i.e., climate). In addition, the aggregated LS of individuals in a country is not necessarily a good proxy for the LS of a community or country as a whole as a community may be influence by factors outside an individual's perception (Cloutier & Pfeiffer, 2015). Aggregation can mask the distribution of LS and is based on an individual perspective and not a community or societal perspective (Kubiszewski, Zakariyya, et al., 2019).

Nevertheless, LS has been shown to correlate well with objective assessments of wellbeing (Oswald & Wu, 2010) and is measured for most UN countries at regular intervals (Helliwell et al., 2019). Therefore, while acknowledging its limitations, we recognize that LS is the best proxy for overall wellbeing for which we have sufficient international data at the moment. We recognize the need for significant additional research on the factors that contribute to sustainable wellbeing and how to measure it.

To highlight these issues, we correlate LS with the 232 official SDG indicators across countries, where data for the indicators exist. We also include additional objective indicators of inequality and income (the Gini coefficient and GDP/per capita) as they are highly relevant to SDGs 8 and 10 and already available for most countries. This paper is an attempt to determine whether a smaller number of SDG indicators can reliably predict LS, an aspect of wellbeing. This process will hopefully help identify key overarching types of indicators that may be more feasible to measure with less resources, but with similar results. Our goal is to provide countries, and policy-makers within those countries, a snapshot of the realities of data availability at present, given the potential cultural variability in LS and limited resources available toward measuring indicators that impact sustainable wellbeing.

2 | METHODS

2.1 | Data

This paper uses the 232 SDG indicators, as published by the UN Statistics Division.[1] These indicators are further subdivided based on gender, age, and location (urban/rural) yielding over 650 indicators. We dropped all indicator where data was available from less than 80 countries. This resulted in an initial subset of 278 indicators. Due to the high number of indicators that were only measured sporadically by a significant number of countries, we used a 20-year average for each indicator by country. We then removed indicators with more than 10% of missing values (104 indicators), ones that were binary or near-binary (6 indicators), and indicators with zero variance (1 indicator)[2] resulting in 167 indicators.

The LS data used in this paper come from the Gallup World Poll (GWP). The GWP uses the "Cantril Ladder" question, asking respondents to imagine a ladder with the best possible life being a 10, and the worse possible life being a 0 (Diener et al., 1985; Gallup, 2009). For comparability, a 20-year average was also taken of LS.

To ensure consistency within our 20-year average results, we also ran the same analyses using data averaged over the last three 5-year periods.

2.2 | Indicator selection

We then proceeded to search for a highly-regulated set of indicators that were predictive of LS by using the least absolute shrinkage and selection operator (LASSO), a constrained linear regression algorithm used in high-dimension data sets (Tibshirani, 1996). The LASSO was calculated using the R package glmnet (Friedman et al., 2010) which is not tolerant of missing values. To enable the reliable use of the LASSO, the R package MICE (van Buuren & Groothuis-Oudshoorn, 2011) was used to impute missing values by using classification and regression trees (CART) (Breiman et al., 1984) derived from the relationships between the data for all countries. Twenty iterations were performed for each set. CART was used because of its robustness and ability to handle multicollinearity. The MICE algorithm draws values sequentially from imputed probability distributions meaning different imputed values are produced each run. Any indicators that had remaining missing values were not used in the LASSO. 144 out of 159 countries had sufficient data to be included in the LASSO analysis.

MICE produced five imputed sets and we applied the LASSO to each set. When using the LASSO, there is a trade-off between model fit (as measured with average mean-squared error) and model regularity (ability of the model to predict unknown values). We selected the most regulated model that was still within one standard error of the best level of model fit for each imputed data set.

2.3 | Model validation and assessment

A 10-fold cross-validation was performed on each of the models from the five imputed data sets to estimate predictive accuracy of the identified indicators. Model selection and cross-validation were both performed in glmnet. Our final model was created using indicators that occurred in all five LASSO models.

Given the complexity of the analysis, we also decided to validate our approach by comparing our results to a simulated data set with no underlying structure. A simulated data set was created with the same number of indicators as our data set with the same pattern of missing indicators. All values were random draws from a normal distribution with mean 0 and variance 1. MICE was again used to impute missing values and a LASSO regression was run. Cross validation was used to assess predictive power and compare to the actual results.

Additional steps were then taken to assess the ability of the identified indicators to assess country performance. First, we ran a multiple regression using the identified indicators to determine model fit as measured by the squared correlation coefficient. We reran this model with GDP per capita and the GINI coefficient added as additional predictor indicators to assess the relationship with these two key indicators, which were missing from the UN SDG indicators.

Next, we used a randomization-based clustering technique to determine the ability of the indicators along with LS to group countries according to their performance in sustainable development. We created a distance matrix for all countries in our data set using Euclidean distance based on standardized measures (z-scores) of the identified indicators with and without LS as an additional indicator. We then used the similarity profile routine (SIMPROF) (Clarke et al., 2008) to determine significant groups of countries at the 5% significance level. SIMPROF is a post-hoc randomization-based algorithm that does not require prior hypotheses to search for and identify multivariate structure (clusters).

2.4 | Developing/transitioning and developed countries

To assess model robustness, the same LASSO regression was run on just the developing/transitioning countries and also the developed countries. Developing and transitioning countries were merged since we are averaging data over 20-years and most of the transitioning countries were developing countries within those 20 years.

3 | RESULTS

3.1 | LASSO regression

The LASSO regression yielded eight indicators that appeared in all five most-regulated models and capture a significant portion of the variability in LS (Table 1). Three out of five models only had these eight indicators, and 10-fold cross validation showed an average mean-squared error of 28%–30%. For comparison, cross validation of our simulated (random) data set showed the best model with 12 variables had an average mean-squared error for simulated LS of 71%.

When these eight indicators are run through an ordinary least squares regression, we find that all indicators are statistically significant, except Manufacturing and IT Use ($N = 144$, as 15 countries

TABLE 1 A list of the indicators that capture a significant portion of the variability in LS in all the countries combined (all – eight indicators), developing and transitioning countries separated out (developing – seven indicators), and the developed countries separated out (developed – five indicators)

SDG	Full name	Abbreviation	All	Developing	Developed
1.4.1	Proportion of population using basic drinking water services, by location (percent)	Water		X	
3.6.1	Death rate due to road traffic injuries (per 100,000 population)	Traffic deaths	X		
3.8.1	Universal health coverage (UHC) service coverage index	Health coverage	X	X	X
3.9.1	Crude death rate attributed to household and ambient air pollution (deaths per 100,000 population)	Indoor & outdoor air pollution	X	X	X
3.9.1	Crude death rate attributed to ambient air pollution (deaths per 100,000 population)	Outdoor air pollution			X
8.4.2 12.2.2	Domestic crop consumption per capita (tonnes)	Crop consumption	X	X	
8.5.2	Unemployment rate (%)	Unemployment	X	X	
9.2.1	Manufacturing value added per capita (dollars)	Manufacturing	X		
12.4.2	Electronic waste generated, per capita (kg)	Electronic waste	X	X	X
17.8.1	Internet users per 100 inhabitants	IT Use	X	X	X

TABLE 2 An ordinary least squares regression for all the countries in the world with LS as the dependent variable and the eight listed indicators as independent indicators

Variable	Parameter	SD	T-STAT	2-tail *p*-value	1-tail *p*-value
(Intercept)	4.373	0.402	1.087e+10	3.60e-20	1.80e-20
Traffic deaths	−0.016**	0.007	−2.198e+00	.030	.015
Health cover	0.021***	0.006	3.679e+00	3.37e-4	1.70e-4
Air pollution	−0.005***	0.001	−3.195e+00	.002	.009
Unemployment	−0.023***	0.006	−3.988e+00	1.09e-4	5.44e-05
Consumption	0.209***	0.053	3.970e+00	1.16e-4	5.82e-05
Manufacturing	2.955e-05	2.685e-05	1.101e+00	.273	.137
Electro. waste	0.047**	0.019	2.491e+00	.014	.007
Internet use	2.739e-4	0.005	5.864e-01	.558	.2793
Multiple R	0.914				
Adjusted R^2	0.836				
F-test (value)	85.77				
F-test (DF numerator)	8				
F-test (DF denominator)	135				
p-value	0				
Residual SD	0.449				
Sum squared residuals	27.19				

***Significant at .01 level based on 2 tail *p* values; **Significant at .05 level; *Significant at .1 level.

were dropped due to missing values). The overall adjusted R-squared was 0.84 (Table 2). These eight indicators span six goals.

When the same regressions were run with 5-year averaged periods, we found results that were consistent with the 20-year average, with an average adjusted R^2 of 0.74.

A 20-year average of GDP per capita ($N = 90$) was added to the eight indicators and the regression was run again. GDP/capita was not significant, and there was no significant change in the overall results. A 20-year average of the Gini coefficient (a measure of income inequality) was also added ($N = 130$) and the regression rerun. The Gini coefficient was significant ($p < .01$) and the sign on the coefficient was positive (i.e., more inequality correlated with higher LS).

It is worth noting that the regression results are not necessarily indicative of the predictive significance of a variable. The eight

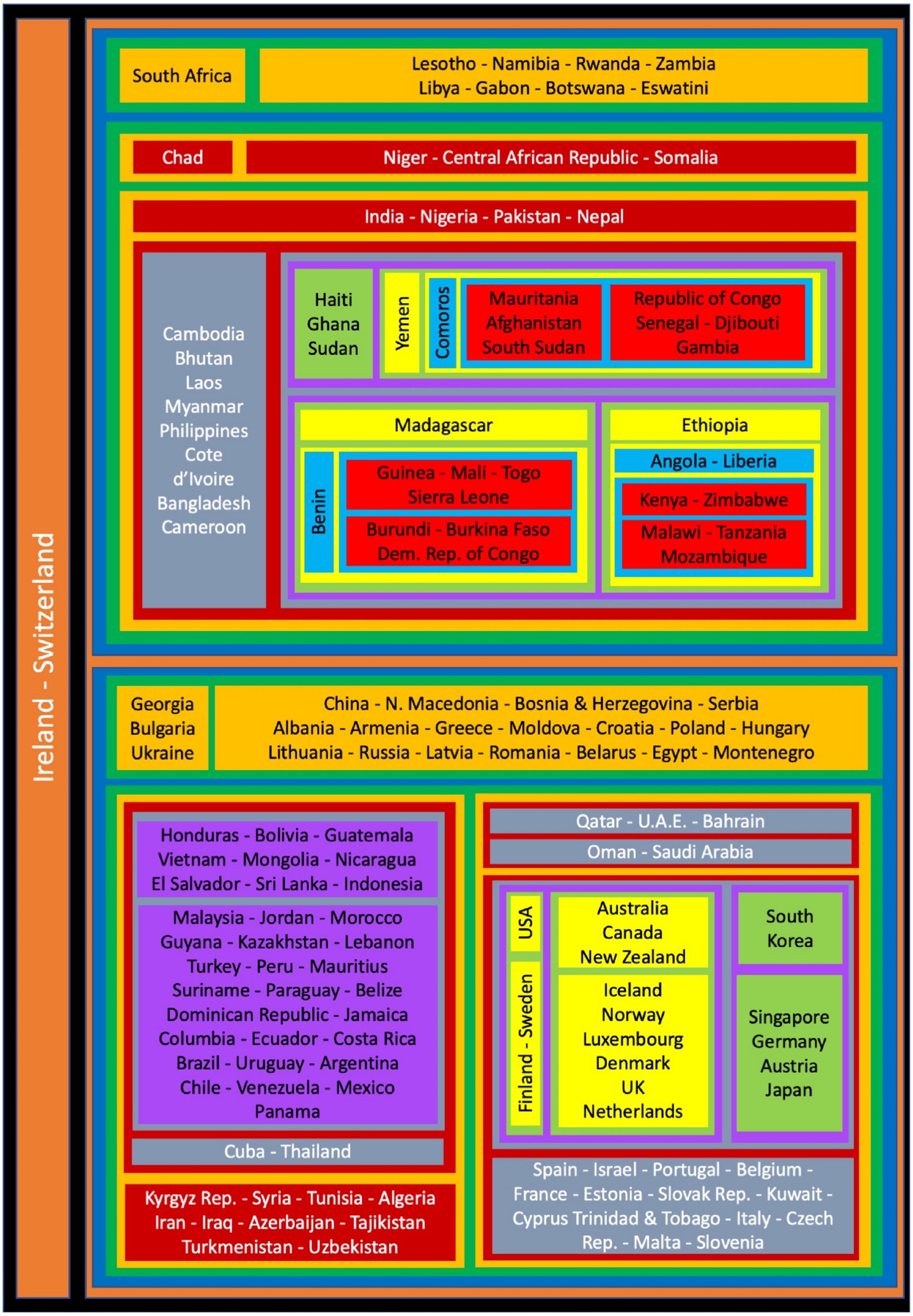

FIGURE 1 Clustering results based on Euclidean distance for the eight standardized indicators plus life satisfaction. Countries in the same box were statistically equivalent based on Euclidean distance. Otherwise, countries are statistically different at the 5% level based on a posthoc randomization test (SIMPROF) (Clarke et al., 2008). When sharing a higher-level box, countries are closer within boxes than to those in lower-level boxes. An alternative way of showing these relationships can be seen in Figure S1. Table S1 show original data [Colour figure can be viewed at wileyonlinelibrary.com]

selected indicators represent the most regulated and cross-validated model, meaning they have been vetted for over-fitting.

The LASSO regression was also run separately for the developing (including transitioning) and developed countries (UN, 2020). Similar indicators were identified in both groups as most correlated with LS. In the developing/transitioning countries, we found that seven indicators closely correlate to LS, while only five indicators were required in the developed countries (Table 1).

3.2 | Clustering

The SIMPROF clustering routine showed a high degree of multivariate structure. The results were based on Euclidean distance for the eight standardized indicators plus life satisfaction. The results were generally the same with and without LS as an additional factor, although clusters were more likely to have statistically significant differences when life satisfaction was included (Figure 1). Countries in the same box were statistically equivalent based on Euclidean distance. Otherwise, countries are statistically different at the 5% level based on a post-hoc randomization test (SIMPROF) (Clarke et al., 2008). Those countries within a higher-level box are closer than those in separate higher level boxes. An alternative way of showing these relationships can be seen in Figure S1 and Table S1 show original data.

4 | DISCUSSION

Eight official SDG indicators can predict 84% of the variation in LS. The fact that it only requires eight indicators to predict 84% of a population's LS, a large aspect of wellbeing, clearly demonstrates that requiring all countries to measure all 232 indicators may be unnecessary and inefficient. Identifying a smaller set of indicators that align more closely with the SDG goals and various aspects of wellbeing, including LS, might provide a more viable and effective strategy.

Our results clearly show that performance in a relatively small number of areas can strongly predict LS. This suggests that sustainable wellbeing can be efficiently and effectively measured with a much smaller number of indicators in the following overarching areas: access to health care, access to food, access to manufactured goods especially electronic and IT goods, access to infrastructure like safe roads and the Internet, access to employment, and access to clean air and water. A small number of indicators would be needed for each of these areas to be effectively assessed. Focusing on a smaller subset of indicators is more effective economically and logistically than trying to track 232 indicators. While further work is needed to assess exactly which subset of indicators and the robustness of the resulting model, our work clearly suggests such a model is possible.

It is also important to remember that identifying these eight indicators does not mean that countries should invest all their resources solely into improving these specific eight indicators (Bevan & Hood, 2006; Espeland & Sauder, 2007). These eight indicators are themselves proxies for larger issues, as are many of the SDG indicators. For example, countries should not attempt to increase the generation of electronic waste just because electronic waste correlates with higher LS. However, they should recognize that the electronic waste indicator is a proxy for access to electronics and manufactured goods. Hence, countries should attempt to increase their population's access to electronics and manufactured goods (Sarriera et al., 2015) while finding ways of reusing or recycling electronic and other waste. If successful, electronic waste will cease to be a good proxy indicator and more direct measures of access to electronics and other manufactured goods will be required.

Likewise, LS increases as the death rate due to road traffic injuries decreases. This indicator is a proxy for a much larger, systemic problem of insufficient or ill-maintained transportation infrastructure and the value of safe infrastructure in general. It implies a need for improved road behavior, including strong law enforcement of traffic laws, within a country (Borowy, 2013; Khorasani-Zavareh et al., 2009; La Torre et al., 2007; WHO, 2017). But it also demonstrates the need for safety across all built infrastructure and better mobility, which may require very different infrastructure. Also, if successful, more direct indicators of safety and mobility will be required.

More direct measures of increasing LS include ensuring universal health coverage, access to food, improving indoor and outdoor air quality, decreasing unemployment, access to electronics and manufactured goods, and providing citizens with reliable infrastructure such as internet access (Adler & Seligman, 2016; Atun et al., 2013; Bartikowski et al., 2018; Çıkrıkci, 2016; Clark et al., 2008; Ferreira et al., 2013; Mee-Udon, 2014; Orru et al., 2016; Welsch, 2006).

GDP per capita is often assumed to be a good proxy for many of these factors. However, GDP per capita has been shown to be significantly correlated with LS in developing countries and less so in developed countries (Easterlin, 2009; Kubiszewski et al., 2013; Proto & Rustichini, 2013). However, GDP was not significant when we added it to the regression with the other eight indicators. This may be due to the fact that the Manufacturing indicator co-varies strongly with GDP. The Manufacturing indicator provides a proxy for how manufactured resources are available to the population and correlates to LS. The pursuit of income, or in this case the Manufacturing indicator, without understanding its relationship to wellbeing, has been shown to have detrimental effects on environmental sustainability and is unrelated to levels of employment (Coscieme et al., 2020). Also, GDP only captures the formal economy. In developing countries, the informal economy is as large, or larger, than the formal economy. This may help explain why globally, GDP per capita was not a significant factor in the model while electronic waste and roadway deaths were.

Although Goal 10 is "Reduced Inequalities," no specific indicator for inequality exists among the 232 official SDG indicators (Fukuda-Parr, 2019; Winkler & Satterthwaite, 2017). We added the Gini coefficient to the regression with our eight indicators and found that it was positively correlated with LS. However, literature has shown that increasing inequality has detrimental effects on an individual's happiness (Kubiszewski, Jarvis, et al., 2019; Wilkinson & Pickett, 2010, 2018). Recently, a happiness-inequality paradox has emerged to explain why the relationship between inequality and LS at the national level may show the opposite result (Barford et al., 2010; Verme, 2011; Wilkinson & Pickett, 2021). This paradox finds that individuals in an unequal society feel like they need to project success and self-reliance by responding to LS surveys in a much more positive way than is necessarily accurate, while individuals in more equal societies feel that stating that they are satisfied with their lives is bragging (Wilkinson & Pickett, 2021).

The Gini coefficient, which represents income inequality within a nation, is only one form of inequality that may be experienced by a population. The percentage of the population with access to clean water is in itself an indicator of inequality, as access to water is essential for life. Access to universal health coverage can also be a means of measuring inequality. The results of our correlation with the Gini coefficient may also be impacted by interactions between these other indicators – further reinforcing need for further research in this important area.

This result raises another issue with over-reliance on cross-country LS as a proxy for sustainable wellbeing. In other words, international comparisons of self-reported LS are problematic due to a range of cultural differences (Graham & Markowitz, 2011), including how individuals respond to inequality. Within countries, lower inequality does lead to better LS. However, comparisons across countries are problematic as they may show opposite effects, as our results illustrate. This is a topic for further research.

Comparing our results with those of a correlation done between an SDG Index and LS (De Neve & Sachs, 2020), we found a stronger relationship between our eight indicators and LS. De Neve and Sachs (2020) found an adjusted R^2 of 0.622 compared to our adjusted R^2 of 0.8356. The De Neve and Sachs (2020) SDG Index synthesizes 63 SDG indicators and adds 14 indicators for OECD countries into an overall assessment (Schmidt-Traub et al., 2017).

Our results can also be compared to a regression of the Human Development Index (HDI) and LS, which has an adjusted R^2 of 0.660 (De Neve & Sachs, 2020). The HDI includes life expectancy, access to education, and GDP per capita. For comparison, we ran an additional model with just GDP per capita and the Gini coefficient as the sole predictor variables. The R^2 value was 0.655 (notably lower than the R2 from our preferred eight indicator models).

All these results show that our eight indicators, as proxies, are strongly correlated with LS. Our eight indicators span the social, environmental and economic aspects of life and provide a holistic perspective on an individual's basic needs. Out of our eight indicators, three relate to SDG Goal 3: “Good Health and Well-being,” while the rest all relate to different goals, including “Goal 8: Decent Work and Economic Growth,” “Goal 9: Industry, Innovation and Infrastructure,” “Goal 12: Responsible Consumption and Production,” and “Goal 17: Partnerships to achieve the Goals.”

When looking at the developing and transitioning countries (separate from the developed countries) we find that seven indicators correlate best with LS. These seven indicators are similar to the eight indicators discussed above except that traffic deaths and manufacturing are no longer part of the list, but water is. Unsurprisingly, these results show that in developing and transitioning countries individuals' LS is most influenced by having their basic needs met, including water, food, employment, clean air, and healthcare, but also access to electronics and the Internet. The individuals living in these countries need more than just income, they require directed goods and services going to those individuals most in need (Stewart, 1979; Streeten, 1979).

Similar indicators were found when looking at only the developed countries. The five indicators that most correlate with LS covered health care, air pollution, and access to electronics and internet. traffic deaths, consumption, unemployment, and manufacturing no longer came out as significant indicators. Air pollution was still significant to individuals' LS, even though most developed countries experience clean air both indoors and outdoors. This shows that this is a critical aspect of individuals' life, whether it is a conscious or unconscious understanding.

Only four indicators occur in all three columns of Table 1: Health Coverage, Indoor and Outdoor Air Pollution, Electronic Waste, and IT Use. This shows the aspects that are universally important to individuals in both developed and developing countries. They cover the three basic elements of sustainability: social, environmental, and economic.

It is also important to remember that LS is completely based on an individual's perception of their own wellbeing. That perception may not completely align with the reality of their own individual wellbeing, or society's sustainable wellbeing. Media and cultural background, for example, play critical roles in creating perceptions (Duffy et al., 2008; Graham, 2011; Graham & Markowitz, 2011) and LS is only a proxy for overall wellbeing. As mentioned above, people are not necessarily well informed about climate (Goal 13), preserving life on land and below water (Goals 14 and 15), inequality (Goal 10), and how the other SDGs relate to their own LS in both the short- and long-term. It is therefore unsurprising that the indicators around these goals are not well correlated with LS.

Differences between countries based on the eight indicators and LS show countries to be strongly clustered according to geography and culture, as well as development status (Figure 1). Countries in the same box were statistically equivalent based on Euclidean distance. Otherwise, countries are statistically different at the 5% level. Those countries within a higher-level box are closer than those in separate higher level boxes. For example, looking at the lower right, the USA is statistically different from all other countries while Finland and Sweden are different from all other counties except each other. Finland and Sweden are statistically closer to the USA than any other countries (the three share the green box). These three together are closer to the nine countries next to them (Australia down to the Netherlands) than all other countries as they share the purple box. At the bottom of the upper blue box, Guinea is closer to Mali than it is to Burundi, closer to Burundi than it is Benin, closer to Benin than it is to Madagascar, and closer to Madagascar than any country not in the green box that they all share.

Switzerland and Ireland stand out from all other countries as they show very high levels of manufacturing but very low levels of deaths from air pollution. The remaining countries split into two more boxes, or groupings; one contains most of the developed countries, while the other most of the developing countries, but not exclusively. These two groupings further split into smaller groupings where regions matter, but a bit less. For example, Cuba and Thailand are grouped together even though they are across the world from each other and do not share a common culture. However, they are at a similar development level and share similar results within our eight indicators. Figure 1 may also show the geography and cultural biases in the way that individuals answer the LS survey and their self-perception. Objective conditions may be similar in different countries, but due to the cultural differences, respondents may claim to have reasonably different LS (Graham & Markowitz, 2011; Kubiszewski, Zakariyya, et al., 2019).

These groupings can be the first step in understanding which countries share which characteristics and problems, so that they can all be addressed in a holistic way. The groupings can also help countries identify other countries with similar problems and dynamics to enable collaboration on solving their mutual problems.

5 | CONCLUSION

The indicators we choose to use, shape the world we create (Kubiszewski et al., 2010). They measure the progress toward our chosen goals (Costanza, Daly, et al., 2016), and shape future policy development. The SDGs have defined societal goals through a set of 17 goals and 232 indicators. These indicators were designed to have a better balance of the three dimensions of sustainable development – social, economic, and environmental – and their governance aspects.

However, indicators are intended to reduce complex, interrelated information to simple scores that are easier to interpret and communicate (Bell & Morse, 2008; Merry, 2011; Morse, 2015; Turnhout et al., 2007). While they are often viewed as objective and direct measures of a concept (Mair et al., 2017), they are actually value-laden, and decisions about them are influenced by politicians, lobbyists, and the media (Morse, 2016; Porter, 1995).

This becomes further complicated when indicators begin to define the concept itself, instead of the other way around (Espeland & Sauder, 2007). This can lead society astray. A good example of this is the Gross Domestic Product (GDP). It was developed as a measure of market activity, but shifted to being a primary measure of societal progress (Fioramonti, 2017; Jackson, 2017). Now, societal progress is viewed as an increase in market activity (Mair et al., 2017) rather than market activity being just one indicator toward the goal of societal progress.

For the SDG indicators to be utilized to their full potential, additional work is still needed to elaborate (1) the complex interconnections between the goals; (2) the means-ends continuum toward an overarching goal; and (3) a "narrative of change" to describe the societal shifts and policy reforms necessary to achieve the SDGs and how this could actually happen within existing socioeconomic and geopolitical circumstances (Costanza, Daly, et al., 2016). Societies need to utilize the SDGs as broad policy goals but also recognize that all attempts to measure progress toward the SDGs and sustainable wellbeing must be taken with a grain of salt. Our eight indicators should not be taken as literal ways of achieving the SDGs. As our previous discussion of electronic waste pointed out, countries should not try to maximize these eight indicators without understanding their connections to LS and sustainable wellbeing.

Our analysis looked at the statistical relationship between the current SDG indicators and average national LS, a proxy for the overarching goal of societal wellbeing. We conclude that most of the current indicators are not necessary and seem to be on the list only because they are, or can be, measured. The current batch of indicators are unable to measure sustainable development holistically, much less sustainable wellbeing. A re-evaluation of the entire process is needed to determine how best to measure progress toward each of the 17 SDGs and how to combine these measures toward sustainable wellbeing (Costanza, Fioramonti, et al., 2016). There is much ongoing research toward this end and we urgently need it to drive the development of systematic policy reforms and societal changes to enable achieving the SDGs at both the national and global level.

ACKNOWLEDGMENTS

We thank our supporting institutions and two anonymous reviewers for their helpful feedback on this manuscript.

ORCID

Ida Kubiszewski https://orcid.org/0000-0003-3264-7899
Diane Jarvis https://orcid.org/0000-0003-4822-8736
Robert Costanza https://orcid.org/0000-0001-6348-8734

ENDNOTES

[1] https://unstats.un.org/sdgs/indicators/database/.

[2] The indicator with zero variance showed whether a population and housing census had been performed in last 10 years. All countries for which we had a response had given an answer of yes, so no variance.

REFERENCES

Adler, A., & Seligman, M. E. P. (2016). Using wellbeing for public policy: Theory, measurement, and recommendations. *International Journal of Wellbeing*, *6*(1), 1–35.

Aksoy, F., & Bayram Arlı, N. (2020). Evaluation of sustainable happiness with Sustainable Development Goals: Structural equation model approach. *Sustainable Development*, *28*(1), 385–392.

Allen, C., Metternicht, G., & Wiedmann, T. (2016). National pathways to the Sustainable Development Goals (SDGs): A comparative review of scenario modelling tools. *Environmental Science & Policy*, *66*, 199–207.

Allen, C., Metternicht, G., & Wiedmann, T. (2017). An iterative framework for national scenario modelling for the Sustainable Development Goals (SDGs). *Sustainable Development*, *25*(5), 372–385.

Ambrey, C. L., Fleming, C. M., & Manning, M. (2014). Perception or reality, what matters most when it comes to crime in your neighbourhood? *Social Indicators Research*, *119*(2), 877–896. https://doi.org/10.1007/s11205-013-0521-6

Atun, R., Aydın, S., Chakraborty, S., Sümer, S., Aran, M., Gürol, I., Nazlıoğlu, S., Ozgülcü, S., Aydoğan, U., Ayar, B., Dilmen, U., & Akdağ, R. (2013). Universal health coverage in Turkey: Enhancement of equity. *The Lancet*, *382*(9886), 65–99.

Badiee, S., Cameron, G., & McNair, D. (2016). *The state of development data funding*.

Bai, X., van der Leeuw, S., O'Brien, K., Berkhout, F., Biermann, F., Brondizio, E. S., Cudennec, C., Dearing, J., Duraiappah, A., Glaser, M., Revkin, A., Steffen, W., & Syvitski, J. (2016). Plausible and desirable futures in the Anthropocene: A new research agenda. *Global Environmental Change*, *39*, 351–362.

Barford, A., Dorling, D., & Pickett, K. (2010). Re-evaluating self-evaluation. A commentary on Jen, Jones, and Johnston (68:4, 2009). *Social Science & Medicine*, *70*, 496–497.

Bartikowski, B., Laroche, M., Jamal, A., & Yang, Z. (2018). The type-of-internet-access digital divide and the well-being of ethnic minority and majority consumers: A multi-country investigation. *Journal of Business Research*, *82*, 373–380.

Bell, S., & Morse, S. (2008). *Sustainability indicators: Measuring the immeasurable?*. Earthscan.

Bevan, G., & Hood, C. (2006). What's measured is what matters: Targets and gaming in the English public health care system. *Public Administration*, *84*(3), 517–538.

Borowy, I. (2013). Road traffic injuries: Social change and development. *Medical History*, *57*(1), 108–138. https://doi.org/10.1017/mdh.2012.83

Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1984). *Classification and regression trees*. Wadsworth Publishing.

Çikrıkci, Ö. (2016). The effect of internet use on well-being: Meta-analysis. *Computers in Human Behavior*, *65*, 560–566.

Clark, A. E., Diener, E., Georgellis, Y., & Lucas, R. E. (2008). Lags and leads in life satisfaction: A test of the baseline hypothesis. *The Economic Journal*, *118*(529), F222–F243.

Clarke, K. R., Somerfield, P. J., & Gorley, R. N. (2008). Testing of null hypotheses in exploratory community analyses: Similarity profiles and biota-environment linkage. *Journal of Experimental Marine Biology and Ecology*, *366*(1), 56–69.

Cloutier, S., & Pfeiffer, D. (2015). Sustainability through happiness: A framework for sustainable development. *Sustainable Development*, *23*(5), 317–327.

Coscieme, L., Mortensen, L. F., Anderson, S., Ward, J., Donohue, I., & Sutton, P. C. (2020). Going beyond Gross Domestic Product as an indicator to bring coherence to the Sustainable Development Goals. *Journal of Cleaner Production*, *248*, 119232.

Costanza, R., Caniglia, E., Fioramonti, L., Kubiszewski, I., Lewis, H., Lovins, H., McGlade, J., Mortensen, L. F., Philipsen, D., Pickett, K., Ragnarsdóttir, K. V., Roberts, D., Sutton, P., Trebeck, K., Wallis, S., Ward, J., Weatherhead, M., & Wilkinson, R. (2018). Toward a sustainable wellbeing economy. *Solutions*, *9*(2). https://thesolutionsjournal.com/2018/04/17/toward-sustainable-wellbeing-economy/

Costanza, R., Daly, L., Fioramonti, L., Giovannini, E., Kubiszewski, I., Mortensen, L. F., Pickett, K. E., Ragnarsdottir, K. V., De Vogli, R., & Wilkinson, R. (2016). Modelling and measuring sustainable wellbeing in connection with the UN Sustainable Development Goals. *Ecological Economics*, *130*, 350–355.

Costanza, R., Fioramonti, L., & Kubiszewski, I. (2016). The UN Sustainable Development Goals and the dynamics of well-being. *Frontiers in Ecology and the Environment*, *14*(2), 59. https://doi.org/10.1002/fee.1231

Costanza, R., Kubiszewski, I., Giovannini, E., Lovins, H., McGlade, J., Pickett, K. E., Ragnarsdóttir, K. V., Roberts, D., De Vogli, R., & Wilkinson, R. (2014). Time to leave GDP behind. *Nature*, *505*(7483), 283–285.

Daly, H. E. (1973). *Toward a steady-state economy*. W.H. Freeman & Company.

Daly, H. E. (1996). *Beyond growth: The economics of sustainable development*. Beacon Press.

De Neve, J.-E., & Sachs, J. D. (2020). Sustainable development and human well-being. In J. F. Helliwell, R. Layard, J. D. Sachs, & J.-E. De Neve (Eds.), *World Happiness Report 2020*. Columbia University.

Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The satisfaction with life scale. *Journal of Personality Assessment*, *49*(1), 71–75. https://doi.org/10.1207/s15327752jpa4901_13

Dolan, P., Layard, R., & Metcalfe, R. (2011). *Measuring subjective well-being for public policy*. London, England.

Duffy, B., Wake, R., Burrows, T., & Bremner, P. (2008). Closing the gaps – Crime and public perceptions. *International Review of Law, Computers & Technology*, *22*(1-2), 17–44. https://doi.org/10.1080/13600860801924899

Easterlin, R. A. (2009). Lost in transition: Life satisfaction on the road to capitalism. *Journal of Economic Behavior & Organization*, *71*(2), 130–145.

Espeland, W. N., & Sauder, M. (2007). Rankings and reactivity: How public measures recreate social worlds. *American Journal of Sociology*, *113*(1), 1–40. https://doi.org/10.1086/517897

Ferreira, S., Akay, A., Brereton, F., Cuñado, J., Martinsson, P., Moro, M., & Ningal, T. F. (2013). Life satisfaction and air quality in Europe. *Ecological Economics*, *88*, 1–10.

Fioramonti, L. (2017). *The world after GDP: Economics, politics and international relations in the post-growth era*. Polity.

Friedman, J., Hastie, T., & Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. *Journal of Statistical Software*, *33*(1), 1–22.

Fukuda-Parr, S. (2019). Keeping out extreme inequality from the SDG agenda – The politics of indicators. *Global Policy*, *10*(S1), 61–69. https://doi.org/10.1111/1758-5899.12602

Gallup. (2009). Understanding how Gallup uses the Cantril scale: Development of the "thriving, struggling, suffering" categories. *Gallup News*.

Giannetti, B. F., Agostinho, F., Almeida, C. M. V. B., Liu, G., Contreras, L. E. V., Vandecasteele, C., Coscieme, L., Sutton, P., & Poveda, C. (2020). Insights on the United Nations Sustainable Development Goals scope: Are they aligned with a 'strong' sustainable development? *Journal of Cleaner Production*, *252*, 119574.

Graham, C. (2011). Adaptation amidst prosperity and adversity: Insights from happiness studies from around the world. *The World Bank Research Observer*, *26*(1), 105–137. https://doi.org/10.1093/wbro/lkq004

Graham, C., & Markowitz, J. (2011). Aspirations and happiness of potential Latin American immigrants. *Journal of Social Research & Policy*, *2*(2), 9–25.

Helliwell, J. F., Layard, R., & Sachs, J. D. (Eds.). (2019). *World Happiness Report 2019*. Sustainable Development Solutions Network.

Helne, T., & Hirvilammi, T. (2015). Wellbeing and sustainability: A relational approach. *Sustainable Development*, *23*(3), 167–175.

Jackson, T. (2017). *Prosperity without growth: Foundations for the economy of tomorrow*. Routledge.

Kapto, S. (2019). Layers of politics and power struggles in the SDG indicators process. *Global Policy*, *10*(S1), 134–136. https://doi.org/10.1111/1758-5899.12630

Khorasani-Zavareh, D., Mohammadi, R., Khankeh, H. R., Laflamme, L., Bikmoradi, A., & Haglund, B. J. A. (2009). The requirements and challenges in preventing of road traffic injury in Iran. A qualitative study. *BMC Public Health*, *9*(1), 486. https://doi.org/10.1186/1471-2458-9-486

Kubiszewski, I., Costanza, R., Franco, C., Lawn, P., Talberth, J., Jackson, T., & Aylmer, C. (2013). Beyond GDP: Measuring and achieving global genuine progress. *Ecological Economics*, *93*, 57–68.

Kubiszewski, I., Farley, J., & Costanza, R. (2010). The production and allocation of information as a good that is enhanced with increased use. *Ecological Economics*, *69*(6), 1344–1354.

Kubiszewski, I., Jarvis, D., & Zakariyya, N. (2019). Spatial variations in contributors to life satisfaction: An Australian case study. *Ecological Economics*, *164*, 106345.

Kubiszewski, I., Zakariyya, N., & Costanza, R. (2018). Objective and subjective indicators of life satisfaction in Australia: How well do people perceive what supports a good life? *Ecological Economics*, *154*, 361–372.

Kubiszewski, I., Zakariyya, N., Costanza, R., & Jarvis, D. (2020). Resilience of self-reported life satisfaction: A case study of who conforms to set-point theory in Australia. *PLoS One*, *15*(8), e0237161. https://doi.org/10.1371/journal.pone.0237161

Kubiszewski, I., Zakariyya, N., & Jarvis, D. (2019). Subjective wellbeing at different spatial scales for individuals satisfied and dissatisfied with life. *PeerJ*, *7*, e6502.

La Torre, G., Van Beeck, E., Quaranta, G., Mannocci, A., & Ricciardi, W. (2007). Determinants of within-country variation in traffic accident mortality in Italy: A geographical analysis. *International Journal of Health Geographics*, *6*(1), 49. https://doi.org/10.1186/1476-072X-6-49

Layard, R. (2006). Happiness and public policy: A challenge to the profession. *The Economic Journal*, *116*(510), C24–C33. https://doi.org/10.1111/j.1468-0297.2006.01073.x

Le Blanc, D. (2015). Towards integration at last? The sustainable development goals as a network of targets. *Sustainable Development*, *23*(3), 176–187.

Lebada, A. M. (2016). *Member states, statisticians address SDG monitoring requirements* [Press release]. http://sdg.iisd.org/news/member-states-statisticians-address-sdg-monitoring-requirements/

MacFeely, S. (2020). Measuring the Sustainable Development Goal indicators: An unprecedented statistical challenge. *Journal of Official Statistics*, *36*(2), 361.

Mair, S., Jones, A., Ward, J., Christie, I., Druckman, A., & Lyon, F. (2017). A critical review of the role of indicators in implementing the sustainable development goals. In W. Leal Filho (Ed.), *Handbook of sustainability science and research* (pp. 41–56). Springer International Publishing.

Mee-Udon, F. (2014). Universal health coverage scheme impact on well-being in rural Thailand. *International Journal of Health Care Quality Assurance*, *27*(6), 456–467. https://doi.org/10.1108/IJHCQA-11-2012-0111

Merry, S. E. (2011). Measuring the world: Indicators, human rights, and global governance: With CA comment by John M. Conley. *Current Anthropology*, *52*(S3), S83–S95. https://doi.org/10.1086/657241

Morse, S. (2015). Developing sustainability indicators and indices. *Sustainable Development*, *23*(2), 84–95.

Morse, S. (2016). Measuring the success of sustainable development indices in terms of reporting by the global press. *Social Indicators Research*, *125*, 359–375.

Nilsson, M., Griggs, D., & Visbeck, M. (2016). Policy: Map the interactions between Sustainable Development Goals. *Nature*, *534*, 320–322.

Orru, K., Orru, H., Maasikmets, M., Hendrikson, R., & Ainsaar, M. (2016). Well-being and environmental quality: Does pollution affect life satisfaction? *Quality of Life Research*, *25*, 699–705.

Oswald, A. J., & Wu, S. (2010). Objective confirmation of subjective measures of human well-being: Evidence from the U.S.A. *Science*, *327*(5965), 576–579. https://doi.org/10.1126/science.1180606

Porter, T. M. (1995). *Trust in numbers: The pursuit of objectivity in science and public life*. Princeton University Press.

Proto, E., & Rustichini, A. (2013). A reassessment of the relationship between GDP and life satisfaction. *PLoS One*, *8*(11), e79358. https://doi.org/10.1371/journal.pone.0079358

Sachs, J. D., Schmidt-Traub, G., Mazzucato, M., Messner, D., Nakicenovic, N., & Rockström, J. (2019). Six transformations to achieve the Sustainable Development Goals. *Nature Sustainability*, *2*(9), 805–814. https://doi.org/10.1038/s41893-019-0352-9

Sarriera, J. C., Casas, F., Bedin, L., Abs, D., Strelhow, M. R., Gross-Manos, D., & Giger, J. (2015). Material resources and children's subjective well-being in eight countries. *Child Indicators Research*, *8*, 199–209.

Schmidt-Traub, G., Kroll, C., Teksoz, K., Durand-Delacre, D., & Sachs, J. D. (2017). National baselines for the Sustainable Development Goals assessed in the SDG Index and Dashboards. *Nature Geoscience*, *10*(8), 547–555. https://doi.org/10.1038/ngeo2985

Soto, C. J. (2015). Is happiness good for your personality? Concurrent and prospective relations of the big five with subjective well-being. *Journal of Personality*, *83*(1), 45–55. https://doi.org/10.1111/jopy.12081

Stewart, F. (1979). Country experience in providing for basic needs. *Finance & Development*, *16*(4), 23–26.

Streeten, P. (1979). From growth to basic needs. *Finance & Development*, *16*(3), 28–31.

Tibshirani, R. (1996). Regression shrinkage and selection via the Lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, *58*(1), 267–288.

Turnhout, E., Hisschemöller, M., & Eijsackers, H. (2007). Ecological indicators: Between the two fires of science and policy. *Ecological Indicators*, *7*(2), 215–228.

UN (2020). World economic situation prospects: Statistical annex. New York, NY

United Nations. (2020). *About the Sustainable Development Goals*. https://www.un.org/sustainabledevelopment/sustainable-development-goals/

van Buuren, S., & Groothuis-Oudshoorn, K. (2011). mice: Multivariate imputation by chained equations in R. *Journal of statistical software*, *45*(i03), 1–67.

Verme, P. (2011). Life satisfaction and income inequality. Washington, DC.

Vorisek, D. & Yu, S. (2020). Understanding the cost of achieving the Sustainable Development Goals. Washington, DC.

Welsch, H. (2006). Environment and happiness: Valuation of air pollution using life satisfaction data. *Ecological Economics*, *58*(4), 801–813.

WHO. (2017). *Road safety*. Copenhagen, Denmark

Wilkinson, R., & Pickett, K. (2010). *The spirit level: Why equality is better for everyone*. Penguin.

Wilkinson, R., & Pickett, K. (2018). *The inner level: How more equal societies reduce stress, restore sanity and improve everyone's well-being*. Penguin.

Wilkinson, R., & Pickett, K. (2021). For better or worse? In C. Deeming (Ed.), *The Struggle for Social Sustainability: Moral Conflicts in Global Social Policy*, (pp. (275–289). Bristol: Bristol University Press.

Winkler, I. T., & Satterthwaite, M. L. (2017). Leaving no one behind? Persistent inequalities in the SDGs. *The International Journal of Human Rights*, (pp. 275–289). Bristol University Press, Bristol. https://doi.org/10.1080/13642987.2017.1348702

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Kubiszewski, I., Mulder, K., Jarvis, D., & Costanza, R. (2021). Toward better measurement of sustainable development and wellbeing: A small number of SDG indicators reliably predict life satisfaction. *Sustainable Development*, 1–10. https://doi.org/10.1002/sd.2234