Journal of Cleaner Production 513 (2025) 145718

Contents lists available at ScienceDirect

Journal of Cleaner Production

journal homepage: www.elsevier.com/locate/jclepro

Mapping the landscape: a survey on the evolution and similarities among alternative well-being indicators

Damiano Meloni *, Maria Giovanna Bosco, Mauro Gallegati

Dept of Management, Università Politecnica Delle Marche, Piazzale Martelli 8 60100 Ancona Italy

A R T I C L E I N F O

JEL classification:
I31
O11
O15

Keywords:
Well-being
Beyond GDP
Text mining
Alternative indicators
Review

A B S T R A C T

The limitations of GDP as a measure of societal well-being have led to a proliferation of "Beyond GDP" indicators, yet systematic comparisons remain scarce. This paper addresses this gap by combining text-mining techniques with empirical analysis to examine similarities and differences among major international well-being indices. Particular attention is given to the treatment of inequality, highlighting its marginal integration and the shortcomings of relying solely on the Gini index. Using Cosine and Jaccard similarity measures, we identify conceptual overlaps, redundancies, and the distinctiveness of environmental indicators such as the Ecological Footprint and the Environmental Performance Index. Empirical validation shows moderate alignment between textual similarity and statistical co-movement, supported by a novel application of spatially stratified heterogeneity analysis, which reveals strong regional structuring of socio-economic, but not environmental, indicators. These findings offer a comprehensive mapping of the well-being indicator landscape and practical guidance for selecting complementary, inequality-sensitive and regionally appropriate metrics for policy and research.

1. Introduction

The critique of gross domestic product (GDP) and the search for alternative indicators have their roots in a long-standing discourse (Fleurbaey and Blanchet, 2013; Stiglitz et al., 2009). Nevertheless, despite ongoing debates, GDP continues to be widely used as a proxy for the well-being of nations, regions, and cities because of its consistent computational standards, which allow for comparability across time and space. However, this accounting convenience has given rise to several criticisms due to its inherent limitations in capturing various facets of human existence that are essential for defining quality of life. For instance, this measure fails to account for socio-economic factors such as employment levels and shows an inverse relationship with indicators of environmental sustainability and broader measures of well-being (Awan and Azam, 2022; Coscieme et al., 2020). A comprehensive measurement of societal progress should require the integration of different scientific schools of thought, bridging subjective well-being, economics, needs theories, capability approaches and ecological perspectives (Jansen et al., 2024).

As a result, there has been a proliferation of alternative indicators of well-being, often referred to as "beyond GDP" measures, in both the academic and policy spheres, which aim to capture a broader range of dimensions, including economic, health, education, environmental sustainability, social capital and happiness. Both scholars and practitioners have proposed different approaches to overcome the limitations of GDP, ranging from single indicators that modify GDP calculations to composite indices and dashboards that incorporate indicators from multiple domains of human life and sustainability. In the academic literature, the social indicator movement has been inspired by frameworks such as the capability approach (Nussbaum, 2000; Sen, 1985) and the literature on happiness and subjective well-being. In addition, the political counterpart proposed alternative measures to GDP, such as the United Nations Development Programme's Human Development Index (Sen and Anand, 1994) and the OECD's Better Life Initiative (OECD, 2011).

Despite the proliferation of well-being indicators, the lack of a universally accepted standard for capturing economic, environmental, and social well-being poses challenges for standardization and comparability. This proliferation of indicators has resulted in a complex landscape, further complicated by disagreements over the ideal approach (e.g., composite indices versus indicator sets) and institutional resistance from national accountants, policymakers and interest groups (van den Bergh, 2009).

While numerous studies have examined individual well-being

* Corresponding author.
E-mail addresses: d.meloni@staff.univpm.it (D. Meloni), m.g.bosco@univpm.it (M.G. Bosco), mauro.gallegati@univpm.it (M. Gallegati).

https://doi.org/10.1016/j.jclepro.2025.145718
Received 26 July 2024; Received in revised form 14 April 2025; Accepted 13 May 2025
Available online 20 May 2025

indices in isolation (e.g. Decancq and Schokkaert, 2016), with some exceptions (e.g., Bleys, 2012; Costanza et al., 2014; Goossens et al., 2007; Offer, 2003), systematic comparisons to identify commonalities and differences remain scarce. In particular, there is no clear understanding of how much two indicators can be considered "competitors" or "complements".

This paper addresses this gap by reviewing the most popular international indicators developed to date and presenting some recent research proposing alternative approaches. Using advanced text-mining techniques, we aim to uncover the underlying structure of existing indices, identify common themes and dimensions in their documentation, discern similarities and dissimilarities among them, and assess how the dimension of inequality, which is increasing at the global level (Chancel and Piketty, 2021), is captured by them. In addition, we aim to contribute to a deeper understanding of the landscape of well-being indicators by combining conceptual analysis with empirical techniques. Specifically, we apply text-mining to uncover similarities between indicators and validate these findings through correlation analyses and spatially stratified heterogeneity (Wang et al., 2016), revealing how well-being indicators behave globally and across regions.

This study has important implications for policymakers, researchers and practitioners invested in measuring and improving societal well-being. By providing insight into the complementarity of well-being indicators and their treatment of inequality, our findings can inform the development of more robust and holistic measures of human well-being. Furthermore, as the importance of incorporating a broader set of well-being indicators is seen as critical in formulating effective government socio-economic strategies (Buettner et al., 2020), policymakers can benefit from our research by gaining insight into the degree of similarity between indices and their intrinsic heterogeneity, which can help in the selection of appropriate measures to avoid redundancy and effectively address inequality.

The paper is structured as follows: in Section 2, we highlight the main limitations of GDP as an indicator of well-being. Section 3 presents recent research on alternative well-being indicators. Section 4 presents the methodological framework and the critical evaluation of the main well-being indicators through a text-mining analysis. In Section 5, we provide empirical validation of the text-mining analysis and a novel application of spatially stratified heterogeneity analysis to well-being indicators. In Section 6, we classify the above indicators according to their consideration of inequality and highlight the principles for ideal indicators. In Section 7, we provide a discussion of the results. The conclusion follows.

2. Conceptual framework

2.1. Inadequacy of GDP as a measure of well-being

Gross domestic product (GDP), the main measure that quantifies the total market value of goods and services produced within a country's borders during a given period, has been widely used to assess a nation's economic well-being. Although some empirical evidence relates GDP to subjective well-being (Li et al., 2022), the comprehensiveness of GDP as a measure of well-being has been questioned due to its inherent limitations. The Stiglitz-Sen-Fitoussi report (Stiglitz et al., 2009), commissioned by the French government, played a key role in highlighting these shortcomings and advocating the consideration of alternative variables. Nevertheless, the need to find indicators that together or separately take into account the different aspects of human well-being has spurred research into alternative, integrated measures. This necessity arose from the difference in meaning between some of the most common terms in economic theory: *growth*, *development*, and *well-being*. Although these terms are used as synonyms in the public debate, they refer to very different economic concepts. While *growth* represents the quantitative increase in income driven by factors such as capital and labor, *development* encompasses a qualitative improvement in productive efficiency alongside income growth. *Well-being*, however, transcends these economic measures, encompassing a broader range of human experience beyond material well-being, inadequate and flawed, as reminded by the Stiglitz–Sen–Fitoussi report.

In contrast to the paradigm shift witnessed in the life sciences towards holistic, system-based approaches, mainstream economic theory remains entrenched in a reductionist framework. While the life sciences emphasize the interconnectedness and interdependence of all biological entities within complex biological systems (e.g. the literature on the ecosystem services), economic theory continues to operate within a framework characterized by the belief that system properties are determined solely by the interactions of its constituent parts, perpetuating the dominance of GDP as the primary indicator of regional well-being.

The consequences of adhering to a GDP-centric paradigm hinder our understanding of the multiple determinants of well-being. Real progress requires moving away from this limited framework and adopting a multidimensional approach that considers economic, social, and environmental dimensions. This paradigm shift requires recognizing the deep interdependencies between these pillars and the inherent complexity of human well-being. (Ciommi et al., 2022).

2.2. The limits of GDP as a comprehensive well-being indicator

The limits of GDP can be traditionally grouped into sets of shortcomings. The first relates to its *economic and financial limits*. GDP fails to consider income distribution and inequalities, despite Oxfam (2024) underlining a concerning trend of income concentration in recent decades. While the wealth of the five richest billionaires in the world has more than doubled in real terms since the beginning of this decade, the poorest 60 % of humanity have seen negligible growth, if any. In the United States, Thomas Piketty and his collaborators found that the incomes of the top 0.001 % of society grew by an average of 6 % per year between 1980 and 2016, while the poorest 5 % saw their incomes fall in real terms.[1] Moreover, several studies still adopt the Gini index as a measure of existing inequalities (UBS, 2023) although it has been proven to be mainly a measure of concentration, not strict inequality (Clementi et al., 2019).

GDP also overlooks the value of non-market goods and services (Stiglitz et al., 2009), such as unpaid labor, volunteer work, and the value of environmental services, and it fails to account for economic activities that are community-based, driven by collaboration and sharing, consequently undervaluing their contribution to overall well-being. Moreover, GDP's reliance on market prices as proxies for value neglects the improving quality of goods and services, as well as differing valuations placed on them by various stakeholders. Furthermore, as an accounting variable, it has a disproportionate focus on what is produced,[2] overlooking the increasing indebtedness of countries and households, treating the replacement of depreciated capital the same as the creation of new capital, and not accounting for future financial liabilities, such as state pensions, thus, providing an incomplete picture of economic health. Paradoxically, GDP indirectly includes the informal

[1] https://www.nytimes.com/interactive/2017/08/07/opinion/leonhardt-income-inequality.html.

[2] GDP accounting has flaws that favor developed economies. From the 1980s onwards, developed economies underwent a significant structural shift. This process accelerated in the 1990s and 2000s. It involved a twofold transformation: a) the outsourcing of many manufacturing activities, and some services, to developing countries, and b) a growing concentration of value-added in high-end service sectors such as banking, real estate and R&D. Notably, this shift has included the production of military weapons systems. Interestingly, the primary growth measure, GDP, has been redefined to better reflect these new areas of specialization in developed economies, to the detriment of developing countries, which have increasingly assumed the role of manufacturing hubs (Assa and Kvangraven, 2021).

economy as a positive: but how does drug trafficking or illegal trade affect welfare? Crime-related activities, social disorder and natural disasters such as forest fires and environmental degradation can all contribute positively to GDP. Expenditure to repair social or environmental damage is included as a positive in the GDP measure, regardless of whether it is a welfare-enhancing or welfare-decreasing activity, while crime-related activities have a multiplier effect on the formal economy (Cobb et al., 1995). For example, defense expenditures such as police, insurance and security only prevent or repair social and environmental costs but are still considered a positive addition to GDP (Leipert, 1989).

Since GDP only records transactions between clearly separated producers and consumers, it underestimates the contribution of new and innovative business models, such as self-production and co-production (Fioramonti et al., 2019).

The second set of shortcomings relates to its *technological limits*. The rise of digital platforms and the gig economy presents challenges to GDP's ability to accurately capture economic activity. Increasing numbers of individuals are engaged in flexible, freelance work, facilitated by digital technologies, using household assets, from computers and smartphones to their homes and cars, yet their contributions often go unaccounted for in traditional GDP calculations.

In addition, the emergence of free digital goods and services further complicates GDP measurement. This is the case of new open-source software, which can substitute for marketed equivalents and has significant economic value despite having no market price, created by the free work of many people (Coyle, 2017). In addition, technological advances such as compression technology and new digital services may not be adequately reflected in GDP figures, leading to an underestimation of their economic impact. For example, new generations of wireless networks are transmitting more data faster than ever before using high-quality compression technologies, while several new services are replacing older ones, such as the replacement of compact discs by streaming music services.

The third set of shortcomings relates to its *social limits*. GDP's scope fails to encompass critical social dimensions of well-being such as the political regime (democratic/authoritarian) and people's participation in voting/elections. Does not consider individual freedom, press freedom, speech freedom, or law compliance, and does not take into account corruption, welfare-reducing opportunistic behaviors and regulatory capture (Dal Bó, 2006). Does not consider intellectual property rights protection. Moreover, GDP overlooks aspects such as labour participation rates, unemployment rates among vulnerable groups, work-life balance, educational attainment, and infrastructure quality, all of which are essential determinants of societal well-being. It does not take into account the distribution of income between individuals, which can enhance personal and social well-being (Wilkinson and Pickett, 2010); it does not take into account social relationships and personal security, health, economic security and longevity (Anheier and Stares, 2002; Michaelson et al., 2009). It ignores cultural differences that lead to different understandings of development goals (Henderson, 1996, 2010). Finally, GDP and welfare in one country can be increased by exporting negative externalities to other countries (Helm et al., 2007).

The fourth set of shortcomings concerns its environmental limitations. GDP's disregard of environmental costs and rates of natural resource depletion is a major drawback. While it may include the cost of cleaning up the environment as a valuable output, it overlooks the long-term negative consequences of exploiting ecosystems. This myopia perpetuates unsustainable economic practices and fails to take into consideration the true costs of environmental degradation. In addition, GDP's failure to recognize the interconnectedness of environmental health and societal well-being means that it inadequately reflects the true state of overall well-being. Table 1 summarizes these limitations.

Table 1
GDP limitations.

Dimension	Limitation
Economic and Financial Limits	Ignores income distribution and inequalities. Excludes non-market goods and services, such as unpaid labor and environmental services. Fails to consider quality improvements in goods and services. Includes informal economic activities without discriminating their impact on well-being. Summarizes perspectives of various entities with differing valuations. Disproportionate focus on production. Fails to account for increasing indebtedness of countries/households. Treats replacement of depreciated capital equally with the creation of new capital. Short-term indicator without considering future financial liabilities. Disregards for informal, community-based, collaborative economic activities. Counts all expenditures as positive, not distinguishing welfare-enhancing from welfare-reducing activities.
Technological Limits	Understates the importance of the digital revolution and contributions of digital platforms. Fails to recognize the equivalence or novelty of digital goods and services. Does not fully capture technological advancements such as compression technology.
Social Limits	Ignores political regime, individual freedoms, press freedom, speech freedom, law compliance. Does not consider corruption, intellectual property rights protection. Neglects labor participation rates, unemployment rates, work-life balance, quality of leisure, etc. Disregards education, infrastructure, energy poverty, health infrastructure, human capital, etc. Excludes economic security, social relations, personal safety, health, and longevity. Overlooks different visions of development goals, cultural differences. Growth in GDP and well-being may come at the expense of ecosystems and workers in developing countries.
Environmental Limits	Omits environmental costs, natural resource depletion rates. Includes costs of environmental remediation as valuable production. Fails to consider long-term negative consequences of short-term exploitation of ecosystems.

Source: Authors' elaboration.

3. Recent research on alternative well-being indicators

On the wave of the above motivations, recent research contributions have addressed the issue of identifying methodologies to capture the interdependencies between economic, social and environmental factors in assessing well-being. The increasing availability of data, both at the macro and micro level, has stimulated researchers to develop new approaches to achieve the dual objective of providing policymakers with simple and concise indicators and capturing the different social, environmental and economic dimensions of well-being and their interrelationships.

We present an overview of the most recent research contributions, highlighting the methodology used, their summary results and, in some cases, the strategy adopted to account for inequality and the estimated correlation with GDP or GDP per capita.

The issue of index correlation has been highlighted by Diener and Suh (1997). If economic indexes show a high correlation with alternative indicators of well-being, the construction of new indexes would simply be redundant. This was the case with the Quality-of-Life Index (Diener, 1995), which is based on Schwartz's (Schwartz, 1992) 45 universal values in two versions: one for poorer nations (Basic QOL Index) and one for wealthier societies (Advanced QOL Index). The Basic QOL Index includes more basic variables - they are usually necessary for life to continue and are often prerequisites for development at the more advanced level. All the basic variables have been standardized and scored so that higher scores represent the desired direction. The correlation with income arises because income accounts for much, but not all, of the variance in the Basic QOL Index. The Advanced QOL Index shows a linear trend with income and is sensitive to differences at the higher income levels. The issue of inequality has been addressed by Boarini et al. (2006) in their study of OECD GDP data. They refer to the concept of equivalized income to consider inequality. Equally distributed income is the level of income that, if distributed equally, would provide society with the same level of welfare as its actual average income in its current unequal distribution. The concept of "equally distributed income" formalizes the possible trade-off between raising the level of total income in society and distributing it more equally at the cost of lower overall efficiency. They find that changes in country rankings of equally distributed income, relative to those based on unadjusted data, depend on the value used for inequality aversion: they are small when using a value of 1 (rank-correlation of 0.98), but higher (rank-correlation of 0.77) when using a coefficient of 10. With a coefficient of 1, the average growth of "equally distributed" income is 1.2 %, compared to 1.4 % for conventional income over the period 1985–2002. With a coefficient of 10, the annual growth rate of "equally distributed" income over the same period falls to 0.6 %. A common debate concerns the presentation of indicators in dashboards or as single indicators. Costanza et al. (2016) proposed a formulation for a hybrid subjective well-being index, a single indicator obtained by weighting the economic, natural, and social dimensions.

In 2020, the Italian BES indicators provided by ISTAT led to further contributions. Ciommi et al. (2020) used the province-level BES indicators for 2004–2016 to obtain 3 composite indicators for each of the main pillars that make up well-being, namely economic, social and environmental. Using a Bayesian latent variable model, they find that economic well-being has deteriorated in almost all provinces, social well-being has improved in almost all provinces, and environmental well-being has increased over time, although the authors show a negative relationship between economic and environmental well-being. Alaimo et al. (2022) use the BES regional-level indicators with a non-aggregative synthesis approach based on Partially Ordered Set Theory (Poset Theory) to provide more "complexity-preserving" insights into well-being. Specifically, they describe each domain of well-being as a partially ordered set and calculate synthetic indicators of regional-level well-being rankings for 2017. They find that the central-north regions have a better well-being situation than southern regions, except for the domain of predatory crime, where the situation is sometimes reversed. This approach allows each region to be assessed in relation to the others, rather than in absolute terms.

Fragkos et al. (2021) apply a CGE model to EU data from the EU Survey on Income and Living Conditions (SILC) and the Household Budget Survey (HBS), which provide relevant data to characterize different households. The aim is to assess the impact of green policies on household income in the EU. They use the Gini index to measure income inequality and the S80/S20 ratio, which provides information on the distribution of income between the richest and poorest 20 % of households. They also calculate the "energy expenditure share of income" indicator, which measures the proportion of households whose energy expenditure as a share of disposable income is more than twice the national median. They find that increased climate policy ambition and determination would almost certainly lead to economic restructuring with potentially regressive distributional effects for the most vulnerable groups. Wang and Chen (2022) use official reports or websites and consider 12 well-known indicators "beyond GDP" and use principal component analysis (PCA) to examine their correlations with GDP per capita (GDPP). They find that most of the indicators studied are not beyond GDP, as they are strongly correlated with GDP and measure similar value dimensions; among the indicators considered in their study, only the Ecological Footprint per capita (EFP), which is strongly negatively correlated with GDP, and the Happy Planet Index (HPI), which has a weak correlation with GDP, are informative and could be considered beyond GDP, as they reveal opposite or different information compared to GDP; they conclude that despite widespread criticism, GDP still plays an important role in economic well-being and social development, as it can explain 65 % of the information in the first pillar of GDP. 61 % of the information in the first principal component (PC) accounts for 51.10 % of the information related to the total of 12 indicators.[3]

Jurčišinová and Štiblárová (2022) use existing macroeconomic data to compute an Adjusted Happy Planet Index (AHPI): they find that the correlation between AHPI and GDP per capita in 2012, 2017 is negative; OLS between AHPI and GDPP is not significant: in spatial models, they run Moran's test to find that the level of AHPI in one country has a positive effect on the level of AHPI in a neighboring country. Barnat et al. (2023) presents an index of inclusive growth based on 21 indicators and built around three pillars - economy, living conditions and inequality - using principal component analysis (PCA). They find that higher levels of inclusive economic growth are generally associated with the more economically developed countries. Castellano et al. (2023) build a novel indicator of efficiency in the production of well-being, including social and environmental costs, based on the OECD Better Life Index; they construct "Our Better Life Index" ("Our BLI") using Data Envelopment Analysis (DEA) to assign an empirical weight to each of the 11 dimensions (sub-indices) of well-being considered in BLI. Their input- and output-oriented VRS-DEA yields five efficient countries: the UK, Iceland, New Zealand, Luxembourg, and the USA; social efficiency is a prerogative of the developed countries, and five countries appear to be both technically and socially efficient: Iceland, Luxembourg, the UK, New Zealand and the US. They also highlight the relationship between GDP and well-being, as measured by the multidimensional indicator "Our BLI". As relative income rises, so does well-being, but this direct relationship breaks down above a certain income threshold. At the micro level, Sebestyén et al. (2024) estimate indicators of urban well-being using multilevel principal component analysis. The study identifies the geometric characteristics of cities and the structure of the natural environment as two main components of urban well-being. For previous extended reviews of alternative well-being indicators, see also Wang and Chen (2022), Pinyol Alberich et al. (2023), Barnat et al. (2023) and, earlier, Giannetti et al. (2015) and Costanza et al. (2014).

3.1. Previous classification of well-being indicators

This section presents a comprehensive overview of various classification schemes that try to provide readers and policymakers with a map to navigate the landscape of well-being indicators. These schemes offer valuable frameworks for understanding the diverse range of measures aimed at capturing different facets of human well-being, economic welfare, and sustainability.

Offer (2003) proposed an origin-based classification scheme, categorizing alternative measures into three classes: extended economic accounts, social indicators, and psychological indicators. This classification highlights efforts to extend traditional economic accounts by

[3] As an indicator of production, GDP gives a high weight to the consumption of resources needed to produce the final output. This leads to a high correlation with other indicators that also give a high weight to consumption as an expression of high well-being. The correlation of GDP - or GDP per capita - with alternative indicators is often not constant, as it decreases above certain thresholds. However, an alternative well-being indicator should be uncorrelated with GDP if it adequately weights the environmental impact and hence the depletion of natural resources, which instead leads to an increase in GDP itself.

incorporating information that comes from different sources and fields. Moreover, it underscores the multidimensionality of well-being beyond purely economic metrics. Similarly, Boarini et al. (2006) and Diener and Suh (1997) followed the same approach to classify the well-being indexes. The former employed a comparable classification scheme, delineating alternative measures into categories such as monetary measures, social conditions, and subjective measures of well-being. The latter relates to the different categories of measures to the philosophical approaches to conceive well-being used by the different research fields.

Taking a complementary view, Goossens et al. (2007) implement an objectives-based approach by categorizing alternative measures according to their objectives with respect to GDP: adjusting GDP, replacing GDP and complementing GDP. This framework highlights the range of approaches used to assess well-being, from the adjustment of traditional economic indicators to the direct measurement of life satisfaction and ecological footprint. Similarly, Keune et al. (2006) schematizes well-being indicators as complementary, corrective and transformative in relation to GDP. Afsa et al. (2008) offers a novel scheme that moves away from traditional statistical measures towards more novel approaches, dividing the indicators into five classes: (i) GDP and other indicators linked to the System of National Accounts (SNA), (ii) dashboards or sets of indicators, (iii) corrected income measures and extended national accounts, (iv) non-monetary aggregated composite indicators, and (v) subjective approaches. Costanza et al. (2009), like Goossens et al. (2007), classify alternative measures into four groups based on their relationship with GDP, whether they correct, replace, complement, or provide composite indices alongside GDP.

Bleys (2012) tries to overcome the limitation of the two previous approaches by proposing the substance-based classification scheme, a new classification scheme based on the underlying concepts they aim to quantify, specifically focusing on well-being, economic welfare, and sustainability. According to Bleys (2012), the classification scheme emphasizes the interconnectedness between well-being, economic welfare, and sustainability, providing a holistic framework for evaluating alternative measures of progress. By categorizing measures based on these underlying concepts, policymakers can better understand the various dimensions of progress and make informed decisions that promote sustainable and inclusive development.

Different from the previous ones, a recent work by Custodio et al. (2023) provides an interesting new approach to classifying the indicators. Instead of providing a direct classification of the well-being indexes, the authors disentangle each indicator and search for studies that use economic, social, and environmental indicators to present an extended framework for monitoring progress in human development. After that, they linked each indicator to the respective SDGs and proposed the extension of social foundation with three additional categories.

Despite differences in approach, all classification schemes emphasize the multidimensional nature of well-being and the need to move beyond purely economic metrics, but they all have some drawbacks. The origin-based approach, which categorizes indicators according to their disciplinary origins, oversimplifies contemporary interdisciplinary research on well-being and may overlook important contributions from fields such as economics. It also tends to neglect indicators of sustainability, limiting its applicability in assessing holistic societal progress. Conversely, the target-based approach places undue emphasis on GDP, despite its acknowledged inadequacy in capturing human or economic well-being. This over-reliance on GDP can distort the assessment of societal progress and prevent the inclusion of broader measures. In addition, both classification schemes suffer from overlap and misclassification, grouping indicators that quantify fundamentally different concepts. For example, indicators such as the Ecological Footprint and the Human Development Index are erroneously categorized as substitutes for GDP in the target-based approach, even though they serve different purposes related to sustainability and well-being, respectively. The introduction of a material-based classification scheme offers a promising alternative by focusing on the underlying concepts that indicators seek to quantify, thus overcoming the limitations of previous approaches. However, the substance-based classification may face operationalization, and interpretation challenges due to the complexity of categorizing measures based on abstract concepts, potentially hindering its widespread adoption and practical utility in policymaking. As a result, the policy relevance of existing classification schemes is limited, as they do not effectively guide policymakers in selecting indicators to monitor societal progress.

In addition to the range of alternative systems of well-being indicators described above, the WISE framework developed by the OECD offers a novel lens for assessing societal progress. By focusing on four interrelated dimensions - well-being, inclusion, sustainability and economy - the WISE approach highlights the need to capture cross-dimensional interdependencies and trade-offs. In particular, it emphasizes that measures of human flourishing cannot be divorced from considerations of distributive justice and environmental resilience, echoing many of the criticisms levelled at GDP for neglecting equity and ecological integrity. The WISE Horizons project operationalizes this framework by synthesizing WISE-oriented indicators into a comprehensive database. These indicators inform and accelerate the shift towards policy agendas that integrate social well-being, environmental stewardship and equitable economic structures - goals shared by many of the “beyond GDP” policies examined in this paper. Incorporating the WISE indicators into existing reviews can strengthen policy relevance by highlighting inequality and sustainability concerns. Moreover, the project’s post-growth perspective provides a reference point for rethinking economic models that put human well-being and planetary health at the centre, thus complementing the existing landscape of Beyond GDP measures.

4. Alternative well-being indicators: similarity analysis

4.1. Methodology and selected indicators

Although previous studies have tried to classify the well-being indices according to different approaches, we believe that a better way to confront them is to check for similarities or dissimilarities, by directly looking at the indicators that compose these indices. To perform this task, we implemented a text-mining analysis that draw from their documentation and uses the names of the indicators that make up the selected indices to compute two similarity measures (i.e. Cosine and Jaccard similarities).[4] Both Cosine and Jaccard similarity provide valuable insights into the similarity of well-being indicators, offering complementary perspectives that can aid in classification and analysis in the field of well-being research. In order to accurately measure the similarity between different well-being indices, we first selected a sample of commonly used indicators for which detailed lists of their individual components were readily available. Table 2 provides an overview of the measures of well-being that have been developed over the years. For our text-mining analysis, we focused specifically on those indices that are composed of discrete, identifiable sub-indicators to ensure that we could fully capture and compare their underlying structures, excluding those indicators that are an algebraic adjustment of GDP.

4.1.1. Text preprocessing

To improve the robustness of our analysis, it is necessary to clean the data through an appropriate text pre-processing phase before calculating the similarity measures. The main idea is to remove unnecessary content concerning the scope of the analysis. We used the following

[4] We decided to include the Jaccard similarity along with the cosine similarity as a robustness check of our results, as they are computed using two different algorithms and express two different concepts of similarity.

Table 2
Lists of well-being indicators.

#	Indicator	Creators	Creation	Field Covered	Type	Goal	Area	Main Limitations
1	Calvert-Henderson Quality of Life Indicators	Calvert Group, Hazel Henderson	2000	Economic, education, environmental, health, and social data	Dashboard	Assess quality of life in the US	USA	Limited to specific dimensions, subjective selection, and weighting of indicators.
2	Canadian Index of Well-being (CIW)	Atkinson Foundation, Canadian Policy Research Networks, and others	2011	Economic, social, environmental, health, and education data	Composite	Measure Canadians' quality of life	Canada	Subjective nature of some indicators. Designed for one Country.
3	Ecological Footprint (EF)	Mathis Wackernagel, William Rees	1996	Ecological footprint data, resource consumption data	Composite	Compare human demand on nature against nature's capacity	Global	It may not fully capture the complexity of ecological impact. Challenges in data collection and measurement.
4	Environmental Performance Index (EPI)	Yale Centre for Environmental Law & Policy, Columbia University Earth Institute	2002	Environmental indicators, pollution data, ecosystem health	Composite	Summarize the state of sustainability worldwide	Global	Subjective weighting of indicators. Data gaps in agriculture, water resources, and biodiversity.
5	Equitable and Sustainable Well-being (BES)	Italian National Institute of Statistics (ISTAT), National Council for Economics and Labour (CNEL)	2010	Economic, social, environmental, health, and education data	Dashboard	Measure well-being in Italy	Italy	Subjectivity in selecting and weighting indicators, data availability and comparability issues.
6	Fulfilment of hierarchical needs index	Matthew Clarke, Sardar M.N. Islam, Sally Paech	2005	Maslow's categories of needs	Composite	Measure the satisfaction of needs across different hierarchical levels	Australia	Subjectivity in defining and measuring needs, cultural differences
7	Gallup-Healthways Well-being Index	Gallup, Healthways	2008	Health, emotional well-being, physical well-being	Composite	Track and understand the well-being of individuals	USA	Limited to the United States, the subjective nature of some indicators is due to self-reported data.
8	Genuine Progress Indicator (GPI)	Redefining Progress	1995	Adjusted GDP, income distribution, environmental factors	Single	Measure sustainable economic welfare	Various countries	Subjectivity in valuing non-market goods and services, and in quantifying certain factors such as social capital and leisure time. Difficult to compare across countries.
9	Genuine Savings Indicator (GSI)	World Bank	1990s	Natural capital, human capital, produced capital, financial capital	Single	Assess the sustainability of economic growth	Global	Challenges in quantifying and valuing natural and human capital.
10	Green GDP	Various countries and organizations	1990s	GDP adjusted for environmental factors	Single	Account for environmental costs in GDP	Various countries	Challenges in the valuation of environmental resources, subjectivity in adjusting GDP for environmental factors
11	Gross National Happiness (GNH)	Government of Bhutan	1972	Happiness indicators, cultural values	Composite	Measure the happiness and well-being of Bhutan's citizens	Bhutan	Cultural differences in defining and measuring happiness, Limit applicability elsewhere.
12	Gross National Income (GNI)	United Nations	–	Income data	Single	Measure the total domestic and foreign output of a country	Global	Same as GDP.
13	Happy Life-Expectancy	Ruut Veenhoven	1996	Public Health and Quality of Life	Composite	Assess the combination of happiness and life expectancy.	Global	Relies on self-reported happiness.
14	Happy Planet Index (HPI)	New Economics Foundation	2006	Life satisfaction, life expectancy, EF	Composite	Measure sustainable well-being	Global	Relies on subjective data for life satisfaction.
15	Human Development Index (HDI)	United Nations Development Programme (UNDP)	1993	Life expectancy, education, income	Composite	Assess achievement in key dimensions of human development.	Global	Does not account for inequality, poverty, or sustainability.
16	Index of Economic Well-being	Centre for the Study of Living Standards	1998	Economic Development and Welfare	Composite	Provide a more holistic view of economic progress beyond GDP	National	Subjectivity in weighting components, difficulty in quantifying non-economic factors
17	Inequality-adjusted Human Development Index (IHDI)	UNDP	2010	Life expectancy, education, income, inequality	Composite	Adjust HDI for inequality	Global	Data limitations, and challenges in quantifying and valuing inequality.
18	Measure of Economic Welfare (MEW)	William Nordhaus, James Tobin	1972	Consumption, leisure, inequality	Composite	Adjust GDP for leisure time and unpaid work.	Global	Subjective valuation of leisure time, and distributional aspects of income and wealth.
19	Net National Income	United Nations	–	Income data	Single	Measure the total income of a nation.	Global	Similar limitations as GNI.

(*continued on next page*)

Table 2 (*continued*)

#	Indicator	Creators	Creation	Field Covered	Type	Goal	Area	Main Limitations
20	Net National Income per capita	United Nations	–	Income data	Single	Measure average income per person.	Global	Similar limitations as GNI and Net National Income.
21	OECD Better Life Index	Organisation for Economic Co-operation and Development (OECD)	2011	Income, housing, jobs, community, education, environment, governance, health, life satisfaction, safety, work-life balance	Dashboard	Allow users to compare well-being across countries	OECD countries	Limited to OECD countries, the subjective nature of some indicators, may not fully capture certain aspects of well-being.
22	OECD Inclusive Growth	OECD	2013	Economic Development and Equity	Composite	Measure economic growth that benefits all segments of society.	OECD countries	Limited coverage of non-economic aspects of inclusivity, data availability
23	Physical Quality of Life Index (PQLI)	Morris David Morris	1978	Life expectancy, infant mortality, literacy	Composite	Measure the quality of life or well-being	Global	Limited coverage, challenges in comparability and measurement.
24	Social Progress Index	Social Progress Imperative	2013	Basic human needs, foundations of well-being, opportunity	Composite	Measure the extent to which countries provide for the social of their citizens.	Global	Subjectivity in selecting and weighting indicators, data availability and comparability issues.
25	Sustainable Society Index (SSI)	Geurt Van de Kerk, Arthur R. Manuel	2010	Personal development, sustainability, and social equity	Composite	Assess the sustainability of a society.	Global	Subjectivity in defining sustainability, and data availability.
26	Weighted Index of Social Progress	Richard J. Estes	1997	Social Development and Equity	Composite	Evaluate social progress.	Global	Subjectivity in weighting indicators, data availability issue.
27	Well-being and Progress Index (WIP)	Luca S. D'Acci	2021	Economic, social, environmental, health, and education data	Composite	Measure well-being and progress.	Global	Subjectivity in selecting and weighting indicators, challenges in data comparability and measurement.
28	World Happiness Report	United Nations Sustainable Development Solutions Network	2012	Various indicators related to happiness and well-being	Composite	Measure subjective well-being and happiness.	Global	Subjectivity in selecting and weighting indicators. Affected by cultural factors.
29	Worldwide Governance Indicators	World Bank	1996	Governance indicators	Dashboard	Measure governance performance.	Global	Subjectivity in selecting and weighting indicators, data comparability issues.

Source: Authors' elaboration.

techniques: tokenization, removal of stop words and unnecessary terms, normalization and stemming.

Tokenization is the process of breaking down a text document into sentences and sentences into words, usually referred to as tokens. As our analysis is based on the similarity of terms, this step is mandatory. Stop words are mainly prepositions and articles that have little or no meaning and are usually removed to reduce noise (e.g. "the", "and", "that"). Normalization helps to remove noise from the distribution of words per topic by converting capitalized words to lowercase, so that "Inequality" and "inequality" appear as the same word. The final step is to find the root of each word to reduce dimensionality without losing generality. We have chosen to use a stemming procedure rather than a lemmatisation one, as this procedure returns the actual root of a term, and we are more concerned with the similarity of the words in terms of meaning.

4.1.2. Cosine similarity

Cosine similarity is a fundamental concept in machine learning, particularly in the domain of natural language processing and information retrieval. It measures the cosine of the angle between two vectors in a multidimensional space, typically representing the frequency of occurrence of terms in a document.

In a nutshell, cosine similarity calculates the cosine of the angle between two vectors, represented as **a** and **b**, in a multi-dimensional space. This is achieved by taking the dot product of the vectors and normalizing it by the product of their magnitudes. Formally, the cosine similarity between **a** and **b** is defined as:

$$\text{C-similarity}(\boldsymbol{a}, \boldsymbol{b}) = \frac{\boldsymbol{a} \cdot \boldsymbol{b}}{|\boldsymbol{a}||\boldsymbol{b}|}$$

Cosine similarity ranges from -1 to 1, with values closer to 1 indicating greater similarity between vectors and values closer to -1 indicating perfect dissimilarity. Orthogonality between the vectors is indicated by a value of 0.

In practical applications, cosine similarity is often employed in document clustering, recommendation systems, and information retrieval tasks. Its computational efficiency and ability to handle high-dimensional data make it an indispensable tool in various machine learning algorithms, facilitating tasks such as document similarity analysis, document clustering, and content-based recommendation systems.

To perform the cosine similarity, after preprocessing, each cleaned and normalized text (representing a well-being indicator) is converted into a TF-IDF vector. TF-IDF vectorization involves the following steps: i) Term Frequency (TF): It measures the frequency of a term (word) in a document. It's computed as the number of times a term appears in a document divided by the total number of terms in the document. This normalization prevents bias towards longer documents; ii) Inverse Document Frequency (IDF): This measures the importance of a term in the whole collection of documents. It's calculated as the logarithm of the ratio of the total number of documents to the number of documents containing the word. Terms that appear frequently across many documents are penalized in IDF; iii) TF-IDF: It combines TF and IDF to produce a weight for each term in each document. The weight reflects the importance of the term in the document relative to the entire collection. It's calculated as TF * IDF.

At the end of the process, each indicator text is transformed into a vector where each element represents the TF-IDF weight of a particular term (word) in the collection of single indicators composed of that index. Then cosine similarities are applied to these vectors.

4.1.3. Jaccard similarity

Jaccard similarity, often utilized in data mining, text analysis, and recommendation systems, quantifies the similarity between two sets by measuring their intersection and dividing by the union of the sets. Thus, it calculates the similarity between two sets by comparing the number of common elements to the total number of distinct elements in the sets. In the context of well-being indicators, Jaccard similarity is valuable for capturing the overlap of specific indicators and it is particularly useful when the focus is on the presence or absence of certain indicators rather than their relative prevalence.

Formally, for sets **A** and **B**, the Jaccard similarity is expressed as:

$$\text{J-similarity}(\boldsymbol{A},\boldsymbol{B}) = \frac{|\boldsymbol{A} \cap \boldsymbol{B}|}{|\boldsymbol{A} \cup \boldsymbol{B}|}$$

Here, $|\mathbf{A} \cap \mathbf{B}|$ denotes the size of the intersection of sets **A** and **B**, and $|\mathbf{A} \cup \mathbf{B}|$ signifies the size of their union.

Jaccard similarity ranges from 0 to 1, where 0 denotes no similarity (no common elements between sets), and 1 signifies complete similarity (identical sets). It is particularly useful when dealing with binary data, categorical data, or text data where the presence or absence of elements is of interest rather than their values. Despite its simplicity, Jaccard similarity offers an effective measure for comparing sets and is especially valuable in scenarios where the size of the sets varies significantly or when dealing with sparse datasets as is our case.

4.2. Text mining result

Fig. 1 shows the results of the cosine (left panel) and Jaccard (right panel) similarity analyses.[5] Each cell represents the degree of similarity between two indicators. A darker shade indicates a higher similarity, suggesting that these indices may measure overlapping aspects of well-being. This can help to identify redundancies or complementarities between these indicators. In addition, these matrices can serve as a tool for integrating different indicators into a unified framework for assessing well-being. By identifying clusters of similar indices, we can develop multidimensional models that encapsulate economic prosperity, social equity, and environmental sustainability.

Several key findings emerge from our comprehensive analysis of the well-being indicators. First, it's worth noting that none of the selected indicators show significant similarity to the environmental indices, namely the Ecological Footprint (EF) and the Environmental Performance Index (EPI). This suggests that these environmental indicators provide a unique set of information that is not captured by other indices. In particular, the EF shows only a low cosine similarity with the Sustainable Society Index (SSI), although this correlation is not supported by the Jaccard metric (C_similarity = 0.082). Surprisingly, the EPI appears to be more closely aligned with other indices than the EF, which has the highest degree of similarity with the SSI. The low similarity observed between the two indices of environmental well-being may be due to the unique calculation methodology of the EF indicator, the implications of which warrant further investigation, making them two complementary indicators of the ecological sphere of well-being.

The second important result is the presence of a "dark triangle" between Happy Life Expectancy, the Happy Planet Index and the Human Development Index (codes 9, 10, 11) in both the left and right panels. This implies an overlap of information between these three indices, suggesting potential redundancy or shared conceptual frameworks that should be taken into consideration when policymakers select their policy indicators.

Third, it's important to note that both the Canadian Well-being Index, the Weighted Social Progress Index and the Equitable and Sustainable Well-being (BES) show the highest degree of similarity to other indicators (as shown in Table 3). This underlines the comprehensive insights that these indices provide, which cannot be fully captured by relying solely on individual measures of well-being.

5. Empirical validation

We present here an empirical exercise to compare the conceptual relationships identified through text-mining with the actual statistical behaviour of several well-being indicators. The aim was to explore whether the definitional overlap identified through linguistic analysis actually corresponds to the way these measures move together in empirical data. Such validation is challenging, mainly because of the inherent divergence between the textual documentation that describes what each indicator is supposed to measure and the real-world datasets that capture how these indicators behave in different contexts. For example, two indicators may use similar language around "quality of life" or "sustainability" but rely on different data sources or weighting schemes, leading to different patterns in practice. Conversely, indicators that are phrased differently may show strong correlations if they capture related socio-economic or environmental phenomena. Additionally, socio-ecological variables may present different distributions in different geographical regions due to an inherent spatial stratified heterogeneity, making indicators comparison across countries questionable (Wang et al., 2016). This section provides a detailed description of the dataset used, how correlations were analysed along with the text-mining result, and how two key statistical tests, the Mantel test and a Spearman rank, were applied, taking into account the possibility of spatial heterogeneity. Despite the insights gained from empirical analysis, a deeper conceptual gap remains. Nevertheless, these efforts represent an attempt to bridge theoretical constructs with empirical reality as far as is currently feasible.

5.1. Data

The textual analysis carried out in section 4 had the advantage of basing the study of the well-being indicators on their documentation regarding the way in which the actual indicator is calculated (e.g. single, composite, dashboard, aggregated form, etc.). In this empirical validation exercise, we were forced to work on a sub-sample of indicators for which an empirical analysis could be carried out: the sub-sample includes those indices represented by a single value and calculated for a large number of countries and years, in order to maintain the generalizability of our previous results. We were therefore forced to exclude single-country indicators, such as the Canadian Index of Well-being and the Equitable and Sustainable Well-being, or indicators for which there was no published time series.

The actual dataset is an unbalanced panel of countries reflecting different levels of data availability for a period from 1970 to 2022. Indicators have been chosen to capture a wide range of social, economic and environmental dimensions, namely: two variants of the Ecological Footprint (consumption-based and production-based), the Environmental Performance Index, the Happy Planet Index, the Human Development Index, the Human Development Index adjusted for inequality, the Physical Quality of Life Index, the Weighted Index of Social Progress and the World Happiness Report. Each of these measures is produced by different organizations or research initiatives, resulting in differences in coverage over time and space. While this imbalance inevitably limits the robustness of certain cross-country comparisons, it does allow approximate correlations to be calculated for each pair of indicators. The overarching aim is to identify which measures tend to rise and fall together across countries and years, thus providing at least a partial indication of whether their underlying constructs overlap in practice. As each indicator is expressed in different units, simple correlation coefficients provide an accessible way to measure pairwise co-movement. The presence of incomplete country-year observations does not erase

[5] Appendix A reports the similarity matrices for both cosine and Jaccard similarity used to present the visual representation in Fig. 1.

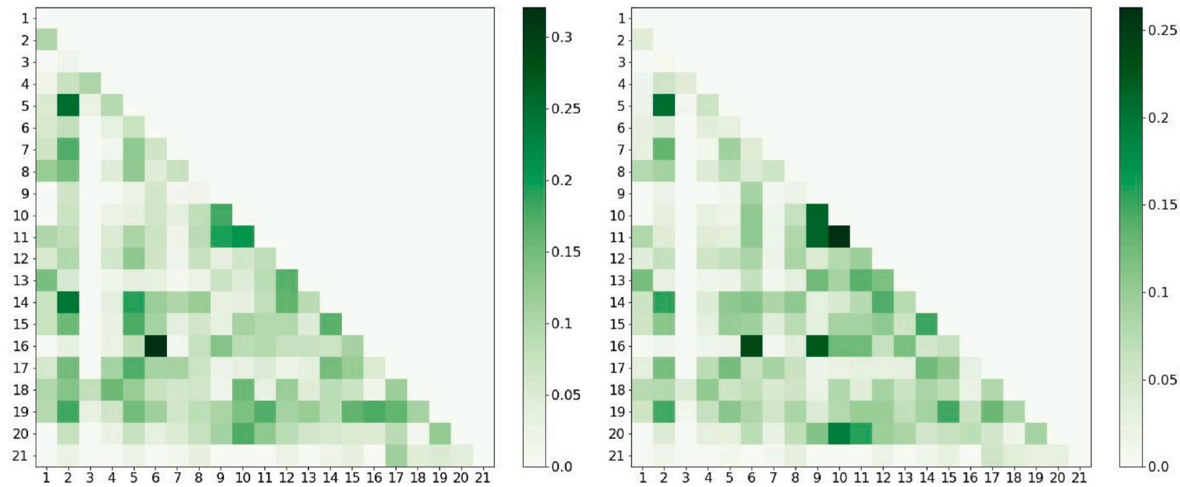

Fig. 1. Cosine (left panel) and Jaccard (right panel) similarity between well-being indicators
Note: 1) Calvert-Henderson Quality of Life Indicators; 2) Canadian Index of Well-being; 3) Ecological Footprint; 4) Environmental Performance Index; 5) Equitable and Sustainable Well-being (BES); 6) Fulfillment of hierarchical needs index; 7) Gallup-Healthways Well-being Index; 8) Gross National Happiness; 9) Happy Life-Expectancy; 10) Happy Planet Index; 11) Human Development Index (HDI); 12) Index of Economic Well-being; 13) Inequality-adjusted Human Development Index (IHDI); 14) OECD Better Life Index; 15) OECD Inclusive Growth; 16) Physical Quality of Life Index; 17) Social Progress Index; 18) Sustainable Society Index (SSI); 19) Weighted Index of Social Progress; 20) World Happiness Report; 21) Worldwide Governance Indicators
The right-hand bar in each panel shows the value of the pairwise similarity scores between pairs of indicators. Values closest to 0 indicate low similarities
Source: Authors' elaboration.

Table 3
Cumulative sum of similarity metrics by well-being indices.

Index	Indicator	Cosine (Sum)	Jaccard (Sum)
1	Calvert-Henderson Quality of Life Indicators	2.103	0.702
2	Canadian Index of Well-being	3.198	1.403
3	Ecological Footprint	1.254	0.098
4	Environmental Performance Index	1.961	0.715
5	Equitable and Sustainable Well-being (BES)	3.060	1.158
6	Fulfilment of hierarchical needs index	2.513	1.375
7	Gallup-Healthways Well-being Index	2.028	0.738
8	Gross National Happiness	2.412	1.148
9	Happy Life-Expectancy	2.095	1.203
10	Happy Planet Index	2.527	1.516
11	Human Development Index (HDI)	2.723	1.700
12	Index of Economic Well-being	2.520	1.358
13	Inequality-adjusted Human Development Index (IHDI)	2.156	1.223
14	OECD Better Life Index	2.906	1.577
15	OECD Inclusive Growth	2.724	1.482
16	Physical Quality of Life Index	2.479	1.256
17	Social Progress Index	2.726	1.237
18	Sustainable Society Index (SSI)	2.594	1.185
19	Weighted Index of Social Progress	3.281	1.596
20	World Happiness Report	2.313	1.261
21	Worldwide Governance Indicators	1.418	0.255

Source: Authors' elaboration.

significant signals where data are available, so correlation results remain valuable for identifying broad patterns.

5.2. Aggregate correlation analysis

Table 4 shows the correlation matrix derived from the entire unbalanced dataset, where each cell contains the correlation coefficient between two indicators across all available country-year observations. Several relationships emerge. For example, the HDI and the IHDI have an extremely high correlation coefficient of 0.98, indicating that these socio-economic measures track each other consistently, despite the IHDI's additional consideration of inequality. In contrast, the Happy Planet Index (HPI) shows a negligible correlation (0.01 and 0.02) with either the consumption-based or the production-based Ecological Footprint variables, highlighting the conceptual divide that separates the HPI's emphasis on subjective well-being and environmental footprint from the more resource-oriented perspective of the EF. Meanwhile, the Environmental Performance Index (EPI) correlates moderately (0.61 and 0.71) with the HDI and IHDI, illustrating that countries with stronger economic or developmental profiles often have higher environmental performance, but not to the extent of the near collinearity seen between human development indicators alone. These observations show that a subset of indicators (i.e., HDI, IHDI, PQLI, WISP and WHR) are tightly clustered, while others, such as HPI and the Ecological Footprint metrics, are more loosely clustered, suggesting that they may

Table 4
Aggregate correlation analysis matrix.

	EF_CONS	EF_PROD	EPI	HPI	HDI	IHDI	PQLI	WISP	WHR
EF_CONS	1								
EF_PROD	0.89	1							
EPI	0.37	0.21	1						
HPI	0.01	0.02	0.37	1					
HDI	0.37	0.31	0.61	0.43	1				
IHDI	0.63	0.42	0.71	0.44	0.98	1			
PQLI	0.53	0.42	0.65	0.61	0.93	0.91	1		
WISP	0.45	0.38	0.42	0.45	0.87	0.82	0.85	1	
WHR	0.61	0.51	0.48	0.58	0.77	0.76	0.66	0.70	1

Source: Authors' elaboration.

measure different dimensions of well-being and sustainability.

5.3. Text-mining similarities

To maintain consistency with the correlation exercise, we performed a restricted text-mining analysis on the subset of indicators for which real-world data had been collected. To maintain consistency, the consumption and production variants of the Ecological Footprint were merged into a single textual descriptor called "EF", as the underlying documentation and concepts are the same. This left us with eight textual entities, mainly EF, EPI, HPI, HDI, IHDI, PQLI, WISP, WHR. As in Section 4, two methods were used to quantify similarities: the cosine approach, which captures vector-based overlaps in word usage and frequency, and the Jaccard index, which focuses on shared versus unique terms. As shown in Tables 5 and 6, pairs such as HDI and IHDI do not exhibit the degree of textual similarity that might be predicted from their near-collinear real-world scores. This discrepancy arises because the IHDI explicitly emphasizes words such as "inequality" or "disparities", whereas the HDI is built around more general references to "income", "education" and "longevity". On the other hand, the EPI and HPI, which share some environmental terminology in their definitions, show moderate textual overlap, although their empirical correlation is modest. These findings underline an important point: textual closeness does not necessarily translate into numerical co-movement, and strong numerical correlation can arise between indicators whose documentation differs significantly in vocabulary and framing.

5.4. Mantel and spearman tests

In order to assess whether the textual similarities were systematically related to the actual correlations, two comparative methods were used. The first was the Mantel test, a method traditionally used in ecology to compare distance matrices. In this case, "distance" is calculated as 1 minus the textual similarity or 1 minus the correlation value, ensuring that large distances correspond to dissimilarities in either semantic or numerical space. The Mantel statistic is a measure of the correlation between these distance matrices, computed as a Spearman rank correlation, and the test relies on permutation to determine significance. For the current dataset, Mantel's r-statistics ranged from approximately 0.55 to 0.86, with statistically significant p-values for all comparisons (see Table 7, world's results). In particular, the comparison between cosine and Jaccard distances yielded $r = 0.864$ with $p = 0.004$, indicating that the two text-mining methods produce broadly consistent patterns of dissimilarity. Meanwhile, Cosine vs. Corr ($r = 0.553$, $p = 0.045$) and Jaccard vs. Corr ($r = 0.568$, $p = 0.053$) suggest that textual and correlation-based distances are in statistically significant, though not strong, agreement.

Secondly, as a robustness test, the Spearman rank correlations were calculated on the flattened top triangles of the matrices. This approach involves listing the pairwise values from each matrix and directly assessing the rank correlation. This simpler test also confirmed a statistically significant relationship ($p < 0.01$ in all cases), with correlation coefficients again around 0.55–0.57 when comparing text-based similarities with the real-world correlation. There is therefore a moderate degree of shared structure between the textual and numerical relationships between the indicators.

Combined, these two tests approximate the empirical validation sought in this section.

5.5. Spatially stratified heterogeneity and empirical validation

This section introduces spatially stratified heterogeneity (SSH), a statistical measure that captures the extent to which indicator values vary systematically across predefined spatial units or strata (Wang et al., 2016). Unlike global measures of variation, SSH focuses on whether and how the distribution of a variable differs between different geographical regions (e.g. continents), highlighting whether certain indicators are influenced by spatial context. For example, if a well-being index is consistently higher or lower in certain regions, this suggests spatial structuring, which SSH can formally detect and quantify. To assess SSH, we follow the methodology developed by Wang et al. (2016), who propose the q-statistic, a ratio that expresses the proportion of the total variance in the data that is due to differences between strata. A higher q-statistic indicates a greater role of spatial structure in explaining the variance of the indicator. In our study, we apply this framework across continents to determine whether well-being indicators exhibit significant regional patterning and to explore the implications of such spatial structuring for measurement and policy. The q-statistic is calculated as follows:

$$\mathrm{q} = 1 - \frac{SSW}{SST} \tag{1}$$

where SST represents the total sum of squares across all observations, and SSW is the within-strata sum of squares. To statistically validate SSH, we employ the F-test based on the noncentral F-distribution proposed by Wang et al. (2016):

$$F = \frac{(N-L)}{(L-1)} \times \frac{q}{1-q} \tag{2}$$

where N is the number of observations, and L represents the number of spatial strata. The significance of SSH is further assessed through the non-centrality parameter λ of the F-statistic:

$$\lambda = \frac{1}{\sigma^2}\left[\sum_{h=1}^{L}\sigma_h^2 - \frac{1}{N}\left(\sum_{h=1}^{L}\sqrt{N_h}\,\mu_h\right)^2\right] \tag{3}$$

A significant SSH is confirmed when $F > F_\alpha$ or when the p-value is below a predefined significance threshold (e.g., 0.1).

We grouped the countries into the five continents and checked for any issues in the data, subsequently we dropped the countries belonging to Oceania due to the high number of missing values encountered. Thus, we stratified the countries into four strata: Europe, Asia, Africa, America.

The results of the SSH analysis (Table 8) reveal considerable spatial heterogeneity for most beyond-GDP indicators among continents. The Ecological Footprint Consumption per Capita (*efconspercap*) and

Table 5
Cosine similarity matrix.

	EF	EPI	HPI	HDI	IHDI	PQLI	WISP	WHR
EF	1							
EPI	0.089	1						
HPI	0.000	0.028	1					
HDI	0.000	0.054	0.217	1				
IHDI	0.000	0.016	0.059	0.128	1			
PQLI	0.000	0.036	0.103	0.109	0.091	1		
WISP	0.021	0.070	0.135	0.193	0.122	0.188	1	
WHR	0.000	0.028	0.170	0.134	0.066	0.064	0.110	1

Source: Authors' elaboration.

Table 6
Jaccard similarity matrix.

	EF	EPI	HPI	HDI	IHDI	PQLI	WISP	WHR
EF	1							
EPI	0.038	1						
HPI	0.000	0.024	1					
HDI	0.000	0.036	0.263	1				
IHDI	0.000	0.012	0.087	0.136	1			
PQLI	0.000	0.013	0.125	0.125	0.118	1		
WISP	0.009	0.061	0.077	0.098	0.066	0.060	1	
WHR	0.000	0.031	0.194	0.156	0.086	0.069	0.087	1

Source: Authors' elaboration.

Table 7
Mantel test and Spearman rank correlation results.

Region	Comparison	Mantel rho	Mantel p-Value	Spearman rho	Spearman p-Value
–	Cosine vs. Jaccard	0.864	0.004***	0.864	0.000***
World	Cosine vs. Corr	0.553	0.045**	0.553	0.002***
	Jaccard vs. Corr	0.568	0.053*	0.568	0.002***
Europe	Cosine vs. Corr	0.510	0.082*	0.510	0.006***
	Jaccard vs. Corr	0.542	0.061*	0.542	0.003***
Asia	Cosine vs. Corr	0.173	0.564	0.173	0.378
	Jaccard vs. Corr	0.127	0.735	0.127	0.519
America	Cosine vs. Corr	0.296	0.313	0.296	0.126
	Jaccard vs. Corr	0.270	0.399	0.270	0.165
Africa	Cosine vs. Corr	0.575	0.012**	0.575	0.001***
	Jaccard vs. Corr	0.532	0.042**	0.532	0.004***

Note: Mantel rho is Spearman's rank correlation on the pairwise distances. A significant p-value indicates that the distance patterns in the two matrices are more similar than random expectation. Regarding Spearman correlations, each pair of matrices was "flattened" by extracting the upper-triangular (off-diagonal) elements, then computing a standard Spearman correlation on those paired values. A significant p-value suggests that pairs of indicators with high similarity in one matrix also tend to be similarly related in the other.

Table 8
Spatial statistic heterogeneity results.

Indicator	q_statistic	F_statistic	Lambda	p_value
EF_CONS	0.088	4.804	0.924	0.451
EF_PROD	0.085	4.662	0.887	0.473
EPI	0.325	24.11	5.398	0.000***
HPI	0.357	27.77	6.306	0.000***
HDI	0.590	71.84	9.156	0.000***
IHDI	0.682	107.2	6.069	0.000***
PQLI	0.651	93.13	8.933	0.000***
WISP	0.455	41.77	4.507	0.002***
WHR	0.459	42.47	12.14	0.000***

Source: Authors' elaboration.

Ecological Footprint Production per Capita (*efprodpercap*) exhibit the lowest q-statistic values (0.088 and 0.085, respectively), with high p-values (0.451 and 0.473), indicating no significant stratification across continents. This suggests that ecological footprint measures are relatively homogeneous globally, likely due to shared environmental impacts and globalized production-consumption dynamics.

Conversely, indicators such as the Human Development Index (HDI, $q = 0.590$, $p < 0.001$) and its inequality-adjusted version (IHDI, $q = 0.682$, $p < 0.001$) display strong spatial structuring, confirming that well-being levels differ significantly across continents. Similarly, the Physical Quality of Life Index (PQLI, $q = 0.651$) and Happiness Index (WHR, $q = 0.459$, $p < 0.001$) show significant spatial patterns, reinforcing the role of geographic, institutional, and socio-economic factors in shaping well-being outcomes.

According to these findings, we decided to repeat the correlation analysis for every continent considered in order to provide a representation of how spatial heterogeneity shapes these relations and our empirical exercise. The results are reported in Tables 9–12. Furthermore, Table 7 reports the comparison with Mantel's and Spearman's tests between the restricted text-mining matrices and the correlation matrix for each country. The table shows that the similarity holds for Europe and Africa, while for Asia and America fails to be maintained. These results confirm the intrinsic difficulties in attempting to bridge theoretical constructs with empirical reality and further support the presence of SSH in well-being dynamics.

However, from a policy perspective, the results are very informative. The observed spatial heterogeneity suggests that a one-size-fits-all approach to well-being improvement may be inadequate at world level. Indicators such as HDI, WHR, and PQLI, which exhibit strong spatial heterogeneity, require region-specific policies to address localized challenges in human development and social progress. Conversely, the absence of spatial heterogeneity in ecological footprint indicators implies that environmental policies could be applied more globally, focusing on shared planetary responsibilities rather than region-specific interventions.

5.6. Limitations and conceptual differences

Despite the attempts at empirical validation, fundamental limitations arise from the conceptual gap between textual definitions and real-world outcomes. Text-mining is designed to capture the presence, frequency and co-occurrence of terms in the official documentation of each indicator, revealing the thematic domains that shape its construction. The actual data, on the other hand, reflects how these constructions translate into measurable outcomes across different populations and geographies. Two indicators may move strongly together over time, for example, HDI and IHDI ($r = 0.98$), but have relatively modest textual overlap due to the strong emphasis on "inequality" in the IHDI's methodology. Conversely, the HPI and EPI may refer to similar environmental concepts in their descriptions, leading to moderate textual similarity, but yield a lower empirical correlation when applied to real-world settings with different resource use and policy priorities. Such examples illustrate how alignment or divergence can occur at different levels: conceptual alignment without numerical parallel, or numerical parallel without significant conceptual overlap. This underlines the fact that textual similarity indicates a shared set of ideas, domains or approaches, whereas numerical correlation indicates shared patterns in actual country scores. Bridging these two perspectives is inherently difficult. Even co-integration tests or other forms of time-series analysis would only show whether indicators share a common trajectory, rather than confirming that they actually measure identical dimensions.

Table 9
Correlation analysis matrix for Europe.

	EF_CONS	EF_PROD	EPI	HPI	HDI	IHDI	PQLI	WISP	WHR
EF_CONS	1								
EF_PROD	0.907	1							
EPI	0.248	0.155	1						
HPI	−0.216	−0.110	0.173	1					
HDI	0.442	0.281	0.485	0.269	1				
IHDI	0.459	0.354	0.547	0.238	0.959	1			
PQLI	0.511	−0.080	0.115	0.176	0.769	0.641	1		
WISP	0.222	0.176	0.457	0.362	0.675	0.689	0.494	1	
WHR	0.409	0.277	0.375	0.440	0.807	0.813	0.588	0.613	1

Source: Authors' elaboration.

Table 10
Correlation analysis matrix for Asia.

	EF_CONS	EF_PROD	EPI	HPI	HDI	IHDI	PQLI	WISP	WHR
EF_CONS	1								
EF_PROD	0.955	1							
EPI	0.196	0.130	1						
HPI	−0.512	−0.541	0.095	1					
HDI	0.674	0.583	0.358	−0.100	1				
IHDI	0.646	0.512	0.513	0.017	0.972	1			
PQLI	0.591	0.556	0.278	−0.308	0.862	0.862	1		
WISP	0.184	0.137	−0.117	0.100	0.659	0.724	0.692	1	
WHR	0.583	0.531	0.169	0.112	0.684	0.672	0.367	0.527	1

Source: Authors' elaboration.

Table 11
Correlation analysis matrix for America.

	EF_CONS	EF_PROD	EPI	HPI	HDI	IHDI	PQLI	WISP	WHR
EF_CONS	1								
EF_PROD	0.777	1							
EPI	0.263	0.231	1						
HPI	−0.378	−0.431	0.187	1					
HDI	0.527	0.576	0.293	−0.058	1				
IHDI	0.784	0.762	0.441	−0.144	0.970	1			
PQLI	0.387	0.340	0.535	0.596	0.876	0.861	1		
WISP	0.156	0.290	−0.148	0.007	0.666	0.671	0.877	1	
WHR	0.542	0.461	0.149	0.363	0.648	0.626	0.633	0.738	1

Source: Authors' elaboration.

Table 12
Correlation analysis matrix for Africa.

	EF_CONS	EF_PROD	EPI	HPI	HDI	IHDI	PQLI	WISP	WHR
EF_CONS	1								
EF_PROD	0.878	1							
EPI	0.271	0.194	1						
HPI	0.150	−0.022	0.194	1					
HDI	0.474	0.392	0.370	0.496	1				
IHDI	0.517	0.359	0.434	0.562	0.961	1			
PQLI	0.435	0.405	0.658	0.620	0.851	0.728	1		
WISP	0.287	0.187	−0.241	0.435	0.632	0.629	0.768	1	
WHR	0.268	0.170	−0.043	0.622	0.353	0.338	0.391	0.631	1

Source: Authors' elaboration.

Thus, the procedures described here represent a coordinated attempt to link textual documentation with empirical behaviour, but they do not negate the structural differences at the core of each methodology. As a result, the moderate positive relationships revealed provide some reassurance that conceptually related indicators can exhibit parallel trends, while also confirming that text-based analysis alone cannot fully capture the complexity inherent in real-world data.

6. Inequality in the well-being indicators

There is considerable variation in the extent to which indicators incorporate measures of inequality. While some indicators include specific measures of inequality, such as the Gini index (which, as noted above, is strictly speaking a concentration and not a pure measure of inequality) or income distribution ratios, others ignore this critical aspect of societal well-being altogether. As Peterson (2014) points out, alternative indicators should consider inequality (in both non-economic and economic terms) to ensure a fair representation of a country's well-being. Table 4 provides an overview of how economic inequality is taken into account in the selected indicators. Evaluating the inclusion of inequality measures is crucial to understanding the extent to which these indicators capture the evolving inequalities within societies, particularly as recent evidence shows that despite a decline in inequality between countries in recent decades, inequality within countries has started to rise again (Chancel and Piketty, 2021). Increasing inequality emerges in both income and wealth distribution and seems to be a process related to capitalism persistence and evolution (Ardeni and Gallegati, 2024; Cetrulo et al., 2024).

In particular, Table 13 shows that only *one-third* of the indices considered take inequality into account in their calculation. This is an important shortcoming in social indicator research, given that inequality is an increasingly crucial aspect of well-being. Indicators that include measures of inequality tend to provide a more comprehensive view of societal well-being by acknowledging inequalities within populations. Conversely, indicators that do not take inequality into account may offer a narrower perspective, focusing only on specific aspects of well-being, such as happiness or environmental sustainability. As Peterson (2014) points out, "a country that ranks high on an average well-being indicator does not necessarily represent an equitably developed society" (Peterson, 2014, p. 591).

A second point to highlight is that the majority of economic inequality indices use the Gini index measure or a transformation of it. Although the Gini index is the most widely used measure of inequality, it has several limitations that don't make it suitable for accounting for inequality in a well-being index. First, it places too much emphasis on the middle of the distribution, whereas other measures, such as the Atkinson family measure (Atkinson, 1970), place more emphasis on the lower end of the distribution. Second, it is a measure of concentration and does not contemplate asymmetries in the Lorentz curve, which could benefit the poorest or the richest households, depending on the direction of the asymmetry, as the Zanardi index does (Clementi et al., 2019; Zanardi, 1964, 1965). Using measures of inequality other than the Gini index would "give greater weight to the deprivation of the deprivation of the poor than of the average (middle class)" (Peterson, 2014, p. 593).

Table 13
Well-being indicators that take inequality into account.

Index	Indicator	Inequality Measure	Description
1	Calvert-Henderson Quality of Life Indicators	NO	
2	Canadian Index of Well-being	YES	Include the Gini index as an indicator.
3	Ecological Footprint	NO	
4	Environmental Performance Index	NO	
5	Equitable and Sustainable Well-being	YES	Disposable Income inequality is computed as the S80/S20 ratio.
6	Fulfilment of hierarchical needs index	NO	
7	Gallup-Healthways Well-being Index	NO	
8	Gross National Happiness	NO	
9	Happy Life-Expectancy	NO	
10	Happy Planet Index	NO	
11	Human Development Index	NO	
12	Index of Economic Well-being	YES	A Gini-based measure of income inequality: Scaled Overall equality = (Poverty Rate *Poverty gap) * 0.75 + Scaled Gini Coeff. (income after tax) * 0.25.
13	Inequality-adjusted Human Development Index	YES	Draws on the Atkinson (1970) family of inequality measures to compute an inequality correction of the HDI.
14	OECD Better Life Index	NO	
15	OECD Inclusive Growth	YES	Adopts real household income for groups of households and introduces an inequality adjustment capturing the difference in income between a given group and average income.
16	Physical Quality of Life Index	NO	
17	Social Progress Index	NO	
18	Sustainable Society Index	YES	Adopts a transformation of the Gini Index: $F(X) = -9/100$*Gini index + 10.
19	Weighted Index of Social Progress	YES	Include the Gini index score.
20	World Happiness Report	NO	
21	Worldwide Governance Indicators	NO	

Source: Authors' elaboration

6.1. Rethinking inequality measures in well-being indices

While our analysis shows that about one third of the well-being indices considered incorporate some measure of inequality - most commonly through the Gini index - this approach may not be sufficient to capture the multifaceted nature of socio-economic disparities. As discussed, the Gini index is primarily a measure of concentration, giving equal weight to all segments of the income distribution and greater sensitivity to changes around the middle of that distribution. This can obscure critical details about the tails of the distribution, where poverty and extreme wealth are found. In addition, the Gini index cannot distinguish the direction of distributional change; for example, it cannot distinguish between scenarios in which inequality is created by an increasingly wealthy minority or by the incomes of a marginalised group falling further behind.

Given the importance of providing policymakers and researchers with a more accurate understanding of inequality, it is important to consider alternative and complementary measures. For example, the Atkinson index (Atkinson, 1970) introduces an explicit parameter for inequality aversion, allowing analysts to give greater weight to changes at the lower end of the distribution. By varying the inequality aversion parameter, the Atkinson index can show how much society values reductions in income differences, especially those affecting the poorest segments. This flexibility provides policy makers with a tool that is consistent with normative judgments about equity and fairness. Another class of indices belongs to the generalised entropy (GE) class, such as Theil's T or Theil's L (Cetrulo et al., 2024). They capture differences in income distributions across the whole range, providing a decomposable structure that allows inequality within and between subgroups to be examined separately. This decomposition capability is particularly relevant for policy makers trying to target interventions to specific regions, socio-economic groups or demographic categories.

Recent literature (Clementi et al., 2019) highlights the importance of considering the asymmetry of the Lorenz curve. The Zanardi index (Zanardi, 1964, 1965) provides information on whether inequality predominantly disadvantages the poorer or the richer segments of society. This directionality can be crucial for understanding how specific policy changes, such as taxation or social transfers, affect different ends of the income spectrum. Including such an index would clarify whether changes in overall inequality reflect, for example, a shrinking middle class or an increase in both extreme poverty and extreme wealth.

To complement traditional measures, indicators that focus on the lower and upper tails of the distribution can be used. For example, poverty measures (e.g. headcount ratios, poverty gaps or severity indices) directly capture the conditions of the poorest groups, while top income share indicators highlight the concentration of resources among the wealthiest. Such measures, used in conjunction with composite well-being indices, could reveal whether a high average level of well-being masks persistent pockets of deprivation or growing extremes of wealth.

By integrating such measures, future well-being indices can provide a more comprehensive and policy-relevant picture. Beyond simply noting that inequality exists, these alternative measures would shed light on the structure and consequences of inequality, guiding targeted policy responses. For example, an index incorporating the Atkinson measure could guide policies to improve the living standards of the bottom deciles, while a Zanardi-based adjustment could inform whether certain redistributive policies truly benefit the disadvantaged or simply redistribute resources between socio-economic groups without altering the underlying imbalances.

To place these suggestions into practice, policymakers should, according to index developers and statisticians, replace or complement the Gini index with measures that provide directionality and sensitivity to extremes by encouraging stakeholder input to define the degree of inequality aversion in Atkinson's measure, ensuring that the index reflects societal preferences, or use GE measures to understand how inequality breaks down across regions, allowing for more context-specific policy interventions. Moreover, considering the presence of spatial heterogeneity across the world, as shown in Section 5.5, it becomes even more crucial to adopt multiple and context-sensitive measures of inequality. This ensures that well-being assessments reflect regional disparities more accurately. A more detailed discussion of these implications is provided in the Discussion section.

Moving beyond the Gini index and using a richer set of inequality measures would allow well-being indicators to better reflect societal conditions. This approach would be particularly valuable for policymakers seeking to promote truly inclusive growth, where improvements in well-being are broadly shared across the population rather than concentrated in particular groups.

7. Discussion

The results of this study provide important insights into the landscape of well-being indicators and relate to previous research in several ways. First, the use of advanced text-mining techniques has enriched the analysis of well-being indices by going beyond traditional taxonomies and empirical correlation methods. In addition, the combination with the empirical exercise allows for a more comprehensive exploration of both conceptual overlaps and empirical relationships between indicators, as advocated by Wang and Chen (2022), and complements previous attempts to classify indicators along dimensions of sustainability, inclusivity and subjective well-being (e.g. Bleys, 2012; Boarini et al., 2006; Costanza et al., 2014). Furthermore, by considering measures of inequality within these indicators, we complement and update previous work on inequality and well-being indices (Peterson, 2014), providing a new and updated perspective.

The distinction between environmental indicators and the overlap between certain well-being indices highlighted by our analysis is in line with previous studies, which emphasize the need to consider the multidimensionality of well-being and avoid redundancies in the selection of indicators. Our empirical exercise has also highlighted the challenges of comparing two conceptually distinct constructs: text-mining similarity and co-movement of indicators. Each method reveals a different dimension of the relationship between indices, and our findings reinforce their complementarity rather than interchangeability.

A second contribution of our study lies in the analysis of spatially stratified heterogeneity. The results show that certain well-being indicators, i.e. HDI, IHDI, PQLI, WHR and WISP, exhibit significant spatial structuring across continents, suggesting that regional differences in socio-economic, institutional and environmental conditions significantly shape how well-being is captured and expressed. In contrast, ecological indicators such as the Ecological Footprint (based on consumption and production) show little spatial heterogeneity, implying a degree of global uniformity in ecological pressures and resource use patterns.

From an ecological economics perspective, this distinction is crucial. The lack of statistically significant spatial structuring in the Ecological Footprint indicators, both consumption and production based, can be interpreted through the lens of ecological economics: environmental pressures transcend national and continental boundaries, with pollution and resource depletion that often diffuse across borders. The globalized nature of supply chains, embodied emissions and ecological outsourcing means that a country's footprint may be largely generated abroad. For example, a country's EF can be outsourced through imports of resource-intensive goods or embodied carbon. This creates cross-border environmental interdependencies, making that country environmentally "clean" domestically, but causing significant environmental degradation abroad (i.e., Global North demand driving deforestation or emissions in the Global South). This reinforces the notion that environmental pressures are planetary in scale and cannot be fully addressed by geographically limited policies. Consequently, environmental indicators such as the Ecological Footprint require global diagnostics and coordinated governance frameworks. In contrast, indicators of socio-economic well-being show greater regional variation and may require context-specific measurement and policy interventions. For example, improving the IHDI or WHR in sub-Saharan Africa requires addressing structural constraints that are very different from those in Europe or Asia. These findings provide the basis for a deeper reflection on differentiated policy needs.

In addition, SSH has implications for the interpretation of indicator correlations and mapping results. As shown in Tables 9–12, indicators that appear to be closely correlated at the global level may diverge significantly within specific regions. For example, while HDI and WHR are strongly correlated in Europe, their association is weaker in Africa, probably due to differences in institutional capacity and cultural perceptions of well-being. This spatial heterogeneity cautions against overgeneralising global findings and underlines the importance of regional context in assessing indicator relationships. It also highlights the uncertainty and limited transferability of global indicator frameworks.

Another important finding is the limited integration of inequality dimensions across the landscape of well-being indicators. Despite increasing within-country disparities worldwide (Chancel and Piketty, 2021), only a third of the indicators reviewed incorporate measures of inequality, and most rely on the Gini index, a measure with known limitations in capturing tail distributions and the direction of change. As a result, many indicators risk providing overly optimistic assessments of well-being by averaging outcomes and masking internal disparities. In policy contexts, such simplifications can lead to misguided strategies that overlook distributional inequities. A wider use of inequality-sensitive measures, such as the Atkinson or Theil indices, would better capture deprivation, wealth concentration and social polarisation, and be more consistent with the goals of inclusive and sustainable development. Furthermore, the SSH analysis strengthens the case for context-sensitive measures of inequality. Indicators such as the IHDI, WHR and PQLI that include or reflect aspects of inequality show strong spatial structuring across continents, implying that the manifestation and impact of inequality are highly context-dependent. For example, inequality in sub-Saharan Africa is often related to basic access to services, whereas in Europe it may be related to wealth accumulation and opportunity gaps. These differences suggest that a single measure of inequality (e.g. the Gini index) is inadequate for global applications and supports the use of stratified, multidimensional measures of inequality that reflect local socio-economic realities. In addition, these results advocate for combining spatial and distributive analyses in policy design: only through such integrated approaches can well-being policies truly support inclusive and equitable development strategies.

Taken together, these findings reveal the complex, multiscalar nature of well-being and ecological stress. They highlight the need for careful selection and interpretation of indicators, taking into account spatial heterogeneity and patterns of inequality. They also provide guidance on how policy makers and institutional actors can better use well-being indicators in a regionally and ecologically appropriate way.

In the stakeholders' perspective, first, our findings highlight the importance of tailoring alternative well-being indicators to explicitly account for inequality and spatial heterogeneity in variable distributions. Indicators that remain highly correlated with GDP have limited added value and should be reconsidered or abandoned in favor of those that capture distinct social or environmental dimensions. This refinement is particularly urgent as rising inequality in the wake of economic, climate and social crises continues to undermine the resilience of marginalised populations.

In a context of growing inequalities in access to resources and intensifying global competition for strategic materials, it is vital that policymakers remain attuned to the potential social and cultural impacts of economic policies. Failure to consider distributional impacts risks deepening inequalities and exacerbating pressures on vulnerable communities and natural ecosystems.

Second, the results caution against the indiscriminate use of well-being indicators in regional aggregates or global rankings. Analyses that group countries together without taking into account context-specific conditions can obscure critical local issues and misinform policy decisions, ultimately contributing to policy failure or increased inequality. Spatially stratified heterogeneity suggests that targeted, regionally adapted policy frameworks are needed to ensure effective interventions. However, the relative spatial uniformity of ecological indicators points to the need for global governance approaches, while the more diverse socio-economic indicators call for locally based solutions. Policy frameworks must therefore operate at multiple scales.

Third, our textual and empirical analyses underscore the value of jointly adopting environmental and socio-economic indicators that are non-overlapping and complementary. These dimensions often give different signals and need to be balanced in order to design policies that are both economically viable and environmentally sustainable.

Finally, the integration of multiple dimensions of well-being, including more appropriate measures of inequality, should be central to governance strategies aimed at inclusive and sustainable growth. Relying solely on averages or inadequate measures of inequality risks perpetuating blind spots in policy design. By embracing multidimensional, inequality-sensitive and spatially-aware indicators, policymakers can better support equitable development and design strategies that are both socially just and environmentally responsible and pave the way for future research and governance frameworks capable of addressing the interlinked challenges of sustainability, equity and global resilience.

8. Conclusion

While GDP has long served as the dominant measure of economic performance, its limitations as a proxy for human well-being have become increasingly apparent. Its narrow focus on market output neglects crucial dimensions of societal progress, including health, inequality, environmental sustainability and subjective well-being. In response, a wide range of alternative "beyond GDP" indicators have emerged, offering more holistic assessments of societal well-being. However, this proliferation has brought new challenges: inconsistencies in classification, overlap between indicators, unclear relationships between their conceptual foundations and empirical behaviour, and limited integration of inequality-sensitive measures.

To address these gaps, this study provides a novel and systematic assessment of the landscape of well-being indicators by combining text-mining analysis with empirical correlation techniques. Unlike traditional classification methods, our approach directly examines the degree of definitional and statistical similarity between composite and dashboard indicators, providing a richer understanding of how these measures relate to each other, both in theory and in practice. The use of cosine and Jaccard similarity allows us to quantify the conceptual overlap between indicators, while the empirical analysis shows how these indicators move together across countries. Together, these methods provide a dual validation of complementarity or redundancy among indexes, highlighting both coherence and divergence.

Our findings provide several important insights. First, certain indicators, mainly environmental ones such as the Ecological Footprint, stand out from the rest, confirming their unique contribution to the measurement of well-being. These indicators show limited conceptual and empirical overlap with socio-economic indices, suggesting that they provide essential, non-substitutable information about planetary boundaries and ecological stress. Conversely, indicators such as Happy Life Expectancy, the Happy Planet Index and the Human Development Index show considerable overlap, raising questions about their added value when used simultaneously. This redundancy poses risks in policy contexts where apparent comprehensiveness may in fact mask blind spots.

Second, incorporating spatially stratified heterogeneity into the analysis shows that well-being indicators are not universally transferable: their relationships and explanatory power vary significantly across regions. Indicators such as the HDI, IHDI, WHR, PQLI and WISP show strong spatial structuring, suggesting that socio-economic and institutional conditions shape the expression and measurement of well-being in regionally specific ways. In contrast, ecological indicators show weaker spatial dependence, supporting the view that environmental pressures are transboundary. This finding is consistent with ecological economics' emphasis on global interdependence, embodied emissions and transboundary ecological footprints. It implies that environmental metrics need to be globally coordinated, whereas social policies need to be locally tailored.

Another key finding is the limited integration of inequality in most well-being frameworks. Despite widening disparities and the increasing salience of distributional concerns in the aftermath of global crises, only a minority of indicators incorporate measures of inequality, most of which rely on the Gini index, a measure with notable limitations. These risks painting an overly optimistic picture of national well-being by masking inequalities within countries. Our findings reinforce the need to incorporate more sophisticated and context-sensitive inequality measures, such as the Atkinson or Theil indices, to better capture polarisation and deprivation. The SSH analysis supports this argument, showing that inequality takes different forms in different contexts: access to basic services in low-income countries, versus opportunity gaps and wealth concentration in advanced economies.

Taken together, these findings have important implications for researchers and policymakers. From a methodological perspective, they highlight the value of integrated approaches, combining textual, empirical and spatial analysis, to refine our understanding of well-being indicators and their limitations. From a policy perspective, they point to the need for indicator portfolios that are both conceptually diverse and empirically complementary, that take account of regional heterogeneity, and that embed inequality concerns at their core. Indicators that are overly correlated with GDP or redundant with each other add little value and can obscure trade-offs and policy priorities.

Looking ahead, future research should explore dynamic relationships between indicators over time to capture the evolution of well-being and its responsiveness to shocks and policies. Longitudinal studies could shed light on how different dimensions of well-being interact and change, revealing synergies or tensions between economic growth, environmental integrity and social inclusion. In addition, more empirical research on the causal effects of well-being policy interventions would help to validate the usefulness of different indices in real-world governance.

In short, this study advances a framework for understanding and comparing well-being indicators that is conceptually sound, empirically robust and spatially sensitive. By identifying complementarities, redundancies and gaps in the current landscape of well-being metrics, we provide practical tools for indicator selection and contribute to a more coherent and effective use of alternative metrics in shaping equitable and sustainable development.

CRediT authorship contribution statement

Damiano Meloni: Writing – review & editing, Writing – original draft, Methodology, Data curation, Conceptualization. **Maria Giovanna Bosco:** Writing – review & editing, Writing – original draft, Supervision, Data curation, Conceptualization. **Mauro Gallegati:** Writing – review & editing, Visualization, Supervision, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Similarity Matrixes

Table A1
Cosine similarity matrix.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0.101	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	0	0.016	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	0.018	0.073	0.100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5	0.052	0.252	0.025	0.092	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6	0.056	0.079	0	0.032	0.072	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
7	0.068	0.173	0	0.008	0.126	0.064	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8	0.121	0.146	0	0.045	0.124	0.043	0.075	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9	0	0.064	0	0	0.025	0.059	0.007	0.013	0	0	0	0	0	0	0	0	0	0	0	0	0
10	0	0.070	0	0.025	0.035	0.063	0.034	0.082	0.178	0	0	0	0	0	0	0	0	0	0	0	0
11	0.098	0.084	0	0.051	0.102	0.068	0.018	0.089	0.193	0.208	0	0	0	0	0	0	0	0	0	0	0
12	0.056	0.096	0	0.058	0.129	0.067	0.015	0.073	0.036	0.058	0.082	0	0	0	0	0	0	0	0	0	0
13	0.145	0.054	0	0.017	0.038	0.033	0.004	0.021	0.071	0.045	0.086	0.167	0	0	0	0	0	0	0	0	0
14	0.072	0.244	0	0.038	0.191	0.119	0.099	0.120	0.033	0.035	0.080	0.161	0.090	0	0	0	0	0	0	0	0
15	0.072	0.153	0	0.027	0.173	0.105	0.038	0.062	0.031	0.109	0.093	0.094	0.047	0.167	0	0	0	0	0	0	0
16	0	0.030	0	0.026	0.084	0.320	0.008	0.075	0.135	0.085	0.093	0.075	0.075	0.068	0.109	0	0	0	0	0	0
17	0.057	0.148	0	0.107	0.172	0.107	0.109	0.068	0.021	0.029	0.049	0.021	0.036	0.148	0.122	0.052	0	0	0	0	0
18	0.098	0.135	0.082	0.153	0.119	0.074	0.058	0.061	0.016	0.155	0.031	0.118	0.045	0.086	0.073	0.017	0.116	0	0	0	0
19	0.091	0.181	0.031	0.064	0.149	0.113	0.062	0.085	0.104	0.141	0.171	0.105	0.122	0.084	0.163	0.175	0.161	0.107	0	0	0
20	0	0.073	0	0.027	0.076	0.038	0.051	0.073	0.109	0.174	0.127	0.087	0.060	0.056	0.052	0.052	0.089	0.006	0.123	0	0
21	0	0.027	0	0	0.023	0	0.009	0.036	0	0	0	0.022	0	0.018	0.035	0	0.113	0.043	0.051	0.040	0

Note: 1) Calvert-Henderson Quality of Life Indicators; 2) Canadian Index of Well-being; 3) Ecological Footprint; 4) Environmental Performance Index; 5) Equitable and Sustainable Well-being (BES); 6) Fulfillment of hierarchical needs index; 7) Gallup-Healthways Well-being Index; 8) Gross National Happiness; 9) Happy Life-Expectancy; 10) Happy Planet Index; 11) Human Development Index (HDI); 12) Index of Economic Well-being; 13) Inequality-adjusted Human Development Index (IHDI); 14) OECD Better Life Index; 15) OECD Inclusive Growth; 16) Physical Quality of Life Index; 17) Social Progress Index; 18) Sustainable Society Index (SSI); 19) Weighted Index of Social Progress; 20) World Happiness Report; 21) Worldwide Governance Indicators.
Source: Authors' elaboration.

Table A2
Jaccard similarity matrix.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0.037	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	0	0.004	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	0.011	0.054	0.038	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5	0.019	0.206	0.005	0.056	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6	0.029	0.040	0	0.033	0.030	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
7	0.025	0.132	0	0.009	0.092	0.037	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8	0.076	0.089	0	0.040	0.069	0.041	0.056	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9	0	0.017	0	0	0.008	0.087	0.006	0.017	0	0	0	0	0	0	0	0	0	0	0	0	0
10	0	0.029	0	0.024	0.016	0.103	0.013	0.063	0.214	0	0	0	0	0	0	0	0	0	0	0	0
11	0.080	0.038	0	0.036	0.030	0.103	0.013	0.079	0.214	0.263	0	0	0	0	0	0	0	0	0	0	0
12	0.034	0.064	0	0.053	0.066	0.083	0.016	0.086	0.042	0.075	0.096	0	0	0	0	0	0	0	0	0	0
13	0.120	0.025	0	0.012	0.016	0.065	0.006	0.030	0.125	0.087	0.136	0.115	0	0	0	0	0	0	0	0	0
14	0.056	0.156	0	0.039	0.105	0.111	0.082	0.105	0.032	0.043	0.075	0.141	0.074	0	0	0	0	0	0	0	0
15	0.055	0.107	0	0.030	0.099	0.093	0.034	0.073	0.031	0.088	0.088	0.103	0.056	0.151	0	0	0	0	0	0	0
16	0	0.013	0	0.013	0.017	0.238	0.006	0.033	0.222	0.125	0.125	0.063	0.118	0.048	0.063	0	0	0	0	0	0
17	0.030	0.119	0	0.071	0.121	0.060	0.088	0.060	0.016	0.023	0.031	0.025	0.031	0.124	0.102	0.024	0	0	0	0	0
18	0.074	0.077	0.041	0.104	0.055	0.068	0.043	0.053	0.021	0.078	0.038	0.086	0.057	0.084	0.071	0.021	0.079	0	0	0	0
19	0.055	0.146	0.009	0.061	0.108	0.081	0.050	0.083	0.040	0.077	0.098	0.098	0.066	0.088	0.149	0.060	0.128	0.083	0	0	0
20	0	0.040	0	0.031	0.032	0.071	0.023	0.066	0.111	0.194	0.156	0.094	0.086	0.049	0.061	0.069	0.050	0.015	0.087	0	0
21	0	0.012	0	0	0.008	0	0.006	0.029	0	0	0	0.017	0	0.014	0.027	0	0.054	0.036	0.027	0.026	0

Note: 1) Calvert-Henderson Quality of Life Indicators; 2) Canadian Index of Well-being; 3) Ecological Footprint; 4) Environmental Performance Index; 5) Equitable and Sustainable Well-being (BES); 6) Fulfillment of hierarchical needs index; 7) Gallup-Healthways Well-being Index; 8) Gross National Happiness; 9) Happy Life-Expectancy; 10) Happy Planet Index; 11) Human Development Index (HDI); 12) Index of Economic Well-being; 13) Inequality-adjusted Human Development Index (IHDI); 14) OECD Better Life Index; 15) OECD Inclusive Growth; 16) Physical Quality of Life Index; 17) Social Progress Index; 18) Sustainable Society Index (SSI); 19) Weighted Index of Social Progress; 20) World Happiness Report; 21) Worldwide Governance Indicators.
Source: Authors' elaboration.

Data availability

Data will be made available on request.

References

Afsa, C., Blanchet, D., Marcus, V., Pionnier, P.-A., Rioux, L., Mira d'Ercole, M., et al., 2008. Survey of existing approaches to measuring socio-economic progress. Commission on the Measurement of Economic Performance and Social Progress. Paris, France.

Alaimo, L.S., Arcagni, A., Fattore, M., et al., 2022. Measuring equitable and sustainable well-being in Italian regions: the non-aggregative approach. Soc. Indic. Res. 161, 711–733. https://doi.org/10.1007/s11205-020-02388-7.

Anheier, H., Stares, S., 2002. Introducing the global civil society index. In: Glasius, Marlies, Kaldor, Mary, Anheier, Helmut K. (Eds.), Global Civil Society Yearbook 2002. Oxford University Press, Oxford, UK.

Ardeni, P.G., Gallegati, M., 2024. La Trappola Dell'Efficienza. Luiss University Press.

Assa, J., Kvangraven, I.H., 2021. Imputing away the ladder: implications of changes in GDP measurement for convergence debates and the political economy of development. New Polit. Econ. 26 (6), 985–1014. https://doi.org/10.1080/13563467.2020.1865899.

Atkinson, A.B., 1970. On the measurement of inequality. J. Econ. Theor. 2 (3), 244–263. https://doi.org/10.1016/0022-0531(70)90039-6.

Awan, A.M., Azam, M., 2022. Evaluating the impact of GDP per capita on environmental degradation for G-20 economies: does N-shaped environmental kuznets curve exist? Environ. Dev. Sustain. 24, 11103–11126. https://doi.org/10.1007/s10668-021-01899-8.

Barnat, N., MacFeely, S., Cantú, F., Peltola, A., Khazhgerieva, A., Panteleev, A., Ryabtsev, N., 2023. Compiling an inclusive growth index: methodological challenges, considerations and conclusions. Stat. J. IAOS 39, 1–21. https://doi.org/10.3233/SJI-230025.

Bleys, B., 2012. Beyond GDP: classifying alternative measures for progress. Soc. Indic. Res. 109, 355–376. https://doi.org/10.1007/s11205-011-9906-6.

Boarini, R., Johansson, A., Mira d'Ercole, M., 2006. Alternative measures of well-being. OECD Economics Department Working Papers No. 33. OECD Publishing, Paris. https://doi.org/10.1787/713222332167.

Buettner, D., Nelson, T., Veenhoven, R., 2020. Ways to greater happiness: a Delphi study. J. Happiness Stud. 21 (8), 2789–2806. https://doi.org/10.1007/s10902-019-00199-3.

Castellano, R., De Bernardo, G., Punzo, G., 2023. Well-being in OECD countries: an assessment of technical and social efficiency using data envelopment analysis. Int. Rev. Econ. 70, 141–176. https://doi.org/10.1007/s12232-023-00413-y.

Cetrulo, A., Sbardella, A., Virgillito, M.E., 2024. L'economia politica delle classi sociali: lavoro operaio e lavoro manageriale nell'Italia delle diseguaglianze. Econ. e Lavoro 12 (1), 151–168. https://doi.org/10.7384/114173.

Chancel, L., Piketty, T., 2021. Global income inequality, 1820–2020: the persistence and mutation of extreme inequality. J. Eur. Econ. Assoc. 19 (6), 3025–3062. https://doi.org/10.1093/jeea/jvab047.

Ciommi, M., Gigliarano, C., Chelli, F.M., Gallegati, M., 2022. It is the total that does [not] make the sum: nature, economy and society in the equitable and sustainable well-being of the Italian provinces. Soc. Indic. Res. 161, 491–522. https://doi.org/10.1007/s11205-020-02331-w.

Clementi, F., Gallegati, M., Gianmoena, L., Landini, S., Stiglitz, J.E., 2019. Mis-measurement of inequality: a critical reflection and new insights. Journal of Economic Interaction and Coordination 14, 891–921. https://doi.org/10.1007/s11403-019-00257-2.

Cobb, C., Halstead, T., Rowe, J., 1995. If the GDP is up, why is America Down? Atl. Mon. 276, 59e78. https://www.theatlantic.com/magazine/archive/1995/10/if-the-gdp-is-up-why-is-america-down/415605/.

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. J. Clean. Prod. 248, 119232. https://doi.org/10.1016/J.JCLEPRO.2019.119232.

Costanza, R., Hart, M., Posner, S., Talberth, J., 2009. Beyond GDP: the need for new measures of progress. The Pardee Papers No. 4—January 2009. Boston University, Boston, MA.

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, 283–285 available at: http://www.nature.com/news/development-time-to-leave-gdp-behind-1.14499.

Costanza, R., Daly, L., Fioramonti, L., Giovannini, E., Kubiszewski, I., Fogh Mortensen, L., Pickett, K.E., Ragnarsdottir, K.V., De Vogli, R., Wilkinson, R., 2016. Modelling and measuring sustainable well-being in connection with the UN sustainable development goals. Ecol. Econ. 130, 350–355. https://doi.org/10.1016/j.ecolecon.2016.07.009.

Coyle, D., 2017. Rethinking GDP. Finance & Development, March 54 (1), 2017. https://www.imf.org/external/pubs/ft/fandd/2017/03/coyle.htm.

Custodio, H.M., Hadjikakou, M., Bryan, B.A., 2023. A review of socioeconomic indicators of sustainability and wellbeing building on the social foundations framework. Ecol. Econ. 203, 107608.

Dal Bó, E., 2006. Regulatory capture: a review. Oxf. Rev. Econ. Pol. 22 (2), 203–225. https://doi.org/10.1093/oxrep/grj013.

Decancq, K., Schokkaert, E., 2016. Beyond GDP: using equivalent incomes to measure well-being in Europe. Soc. Indic. Res. 126, 21–55. https://doi.org/10.1007/s11205-015-0885-x.

Diener, E., 1995. A value based index for measuring national quality of life. Soc. Indic. Res. 36 (2), 107–127. Oct., 1995. https://www.jstor.org/stable/27522867.

Diener, E., Suh, E., 1997. Measuring quality of life: economic, social, and subjective indicators. Soc. Indic. Res. 40, 189–216.

Fioramonti, L., Coscieme, L., Mortensen, L.F., 2019. From gross domestic product to well-being: how alternative indicators can help connect the new economy with the sustainable development goals. The Anthropocene Review 6 (3), 207–222. https://doi.org/10.1177/2053019619869947.

Fleurbaey, M., Blanchet, D., 2013. Beyond GDP: Measuring Welfare and Assessing Sustainability. Oxford University Press, Oxford UK.

Fragkos, P., Fragkiadakis, K., Sovacool, B., Paroussos, L., Vrontisi, Z., Charalampidis, I., 2021. Equity implications of climate policy: assessing the social and distributional impacts of emission reduction targets in the european union. Energy 237. https://doi.org/10.1016/j.energy.2021.121591.

Giannetti, B.F., Agostinho, F., Almeida, C.M., Huisingh, D., 2015. A review of limitations of GDP and alternative indices to monitor human wellbeing and to manage eco-system functionality. J. Clean. Prod. 87, 11–25. https://doi.org/10.1016/j.jclepro.2014.10.051.

Goossens, Y., Mäkipäa, A., Schepelmann, P., van de Sand, I., Kuhndtand, M., Herrndorf, M., 2007. Alternative Progress Indicators to Gross Domestic Progress (GDP) as a Means Towards Sustainable Development. IP/A/ENVI/ST/2007-10. Brussel, Belgium: Policy Department, Economic and Scientific Policy (European Parliament).

Helm, D., Smale, R., Phillips, J., 2007. Too good to be true? The UK's climate change record. Oxford, UK. http://www.dieterhelm.co.uk/node/656.

Henderson, H., 1996. Measuring progress: not by GNP alone. New Economic Indicators Highlight Quality of Life. http://www.lightparty.com/Visionary/MeasuringProgress.html.

Henderson, H., 2010. Continuing the conversation on the politics of happiness is there a "politics of Happiness". http://csrwiretalkback.tumblr.com/post/720597834/continuing-the-conversation-on-the-politics-of.

Jansen, A., Wang, R., Behrens, P., Hoekstra, R., 2024. Beyond GDP: a review and conceptual framework for measuring sustainable and inclusive wellbeing. Lancet Planet. Health 8 (9), e695–e705. https://doi.org/10.1016/S2542-5196(24)00147-5.

Jurčišinová, V., Štiblárová, L., 2022. On what really matters: evidence from alternative well-being indicator in EU-28 countries. Statistika: Statistics & Economy Journal 102 (1). https://doi.org/10.54694/stat.2021.39.

Keune, L., Elzinga, T., Ruyter, T., 2006. Meta Economische Verkenningen—Voorstel Voor De Ontwikkeling Van Een Duurzame En Solidaire Macro Economische Verkenning ±. Attac/Voor de Verandering/XminY, Amsterdam, The Netherlands.

Leipert, C., 1989. Social costs of the economic process and national accounts: the example of defensive expenditures. J. Interdiscipl. Econ. 3, 27e46. https://doi.org/10.1177/02601079X8900300104.

Li, F., Mu, W., Li, S., Li, X., Zhang, J., Chen, C., Zhou, M.J., 2022. Income and subjective well-being: test of a multilevel moderated mediation model. Applied Research in Quality of Life 17 (4), 2041–2058. https://doi.org/10.1007/s11482-021-10017-9.

Michaelson, J., Abdallah, S., Steuer, N., Thompson, S., Marks, N., 2009. National Accounts of well-being: Bringing Real Wealth onto the Balance Sheet. New Economics Foundation, London.

Nussbaum, M., 2000. Women and Human Development: the Capabilities Approach. Cambridge University Press, Cambridge.

OECD, 2011. How's Life? Measuring well-being. OECD Publishing, Paris. https://doi.org/10.1787/9789264121164-en.

Offer, A., 2003. Economic welfare measures and human well-being. In: David, P., Thomas, M. (Eds.), The Economic Future in Historical Perspective. Oxford University Press, Oxford, UK, pp. 371–399.

Oxfam, 2024. Inequality inc. https://oi-files-d8-prod.s3.eu-west-2.amazonaws.com/s3fs-public/2024-01/Davos%202024%20Report-%20English.pdf.

Peterson, L., 2014. The measurement of non-economic inequality in well-being indices. Soc. Indic. Res. 119, 581–598. https://doi.org/10.1007/s11205-013-0513-6.

Pinyol Alberich, J., Llorente-González, L.J., Ramezankhani, M.J., Bimpizas-Pinis, M., Lowe, B.H., 2023. Using macroeconomic indicators to enact an ambitious circular economy. Circ. Econ. Sustain. 3 (3), 1515–1544. https://doi.org/10.1007/s43615-022-00232-3.

Schwartz, S.H., 1992. Universals in the content and structure of values: theoretical advances and empirical tests in 20 countries. In: Zanna, M. (Ed.), Adv. Exp. Soc. Psychol. 25, 1–65. https://doi.org/10.1016/S0065-2601(08)60281-6.
Sebestyén, V., János Trájer, A., Domokos, E., Torma, A., Abonyi, J., 2024. Objective well-being level (OWL) composite indicator for sustainable and resilient cities. Ecol. Indic. 158, 111460. https://doi.org/10.1016/j.ecolind.2023.111460.
Sen, A., 1985. Commodities and Capabilities. North-Holland, Amsterdam.
Sen, A., Anand, S., 1994. Human Development Index: Methodology and Measurement. New York.
Stiglitz, J.E., Sen, A., Fitoussi, J.-P., 2009. Report by the commission on the measurement of economic performance and social progress. www.stiglitz-sen-fitoussi.fr.
UBS, 2023. Global wealth report 2023. https://www.ubs.com/global/en/family-office-uhnw/reports/global-wealth-report-2023/_jcr_content/pagehead/link2.0577489710.file/PS9jb250ZW50L2RhbS9hc3NldHMvd20vZ2xvYmFsL2ltZy9nbG9iYWwtZmFtaWx5LW9mZmljZS9kb2NzL2d3ci0yMDIzLWVuLTIucGRm/gwr-2023-en-2.pdf.
van den Bergh, J., 2009. The GDP paradox. J. Econ. Psychol. 30, 117–135.
Wang, B., Chen, T., 2022. Social progress beyond GDP: a principal component analysis (PCA) of GDP and twelve alternative indicators. Sustainability 14, 6430. https://doi.org/10.3390/su14116430, 2022.
Wang, J.-F., Zhang, T.-L., Fu, B.-J., 2016. A measure of spatial stratified heterogeneity. Ecol. Indic. 67, 250–256. https://doi.org/10.1016/j.ecolind.2016.02.052.
Wilkinson, R.G., Pickett, K., 2010. The Spirit Level: Why Greater Equality Makes Societies Stronger. Bloomsbury Press, New York.
Zanardi, G., 1964. Della Asimmetria Condizionata Delle Curve Di Concentrazione. Lo Scentramento. Rivista Italiana Di Economia, vol. 18. Demografia e Statistica, pp. 431–466.
Zanardi, G., 1965. L'asimmetria statistica delle curve di concentrazione. Ric. Econ. 19, 355–396.