Discover Global Society

Index of active participation in society

Satyendra Chakrabartty[1]

© The Author(s) 2023   OPEN

Abstract

The concept of Active citizenship is gaining importance. Ignoring the issues of selection of dimensions and items, the paper addresses methodological issues and aggregation of indicators of multidimensional active participation index (PI) by two methods. Policy makers and researchers can take advantages of the proposed methods of arithmetic aggregation (Method-1) with normally distributed transformed scores or multiplicative aggregation (Method-2) without scaling and choosing weights. Thus, the Method-2 avoids the problems of rank robustness specific to selection of weights. Multiplicative aggregation proposed in Method-2 can be converted to additive model by taking logarithms. Normal distributions of two or more groups by Method-1 are likely to give rise to lower value of Ginis coefficient indicating equality. Avoiding major limitations of ordinal scores, both methods satisfy desired properties, analysis under parametric set up for meaningful comparisons including testing of statistical hypothesis, identification of critical dimensions, ranking of the dimensions by elasticity, assessment of progress/decline of PI, etc. Method-2 offers more generalized approach satisfying time reversal test and formation of chain indices. However, test of normality is required for this method unlike the Method-1 which ensures normally distributed scores.

1 Introduction

The concept of Active citizenship (AC) is related to civic engagement and building social capital and is defined as participation in civil society, community and political life, characterized by mutual respect and non-violence and in accordance with human rights and democracy [1].

It deals with issues relating to *rights* and *responsibilities*. Clearly, form and content of AC varies with social values, cultural and political to environmental activities at local, regional, national and even international levels. Thus, it excludes participation in extremist groups promoting intolerance and violence.

In a democratic system, quality of citizenship depends on behavior of community members in terms of active participations in politics, society and community along with values. Improving economy, social, and political effectiveness of a country depends heavily on active participation of its citizens. In fact, the concept of good government differs in different countries. Attempts have been made to assess quality of citizenship and degree of active participation [2] with different number of indicators and associated dimensions. Factors of AC at individual level and also at national level using a multilevel model were identified [3] and found higher AC in countries with high GDP, equal distribution of income and heterogeneous religious climate.

✉ Satyendra Chakrabartty, chakrabarttysatyendra3139@gmail.com | [1]Indian Statistical Institute, Kolkata, India.

Measurement of Participation Index (PI) means choosing a real valued function f from n-dimensional space corresponding to n- number of indicators. Quality of PI depends on properties of such function and measurement procedures adopted for the indicators. Ignoring the issues of selection of dimensions and items, the paper gives two methods of aggregation of discrete indicator scores to continuous scores following normal distribution for better and meaningful uses of PI and satisfying desired properties including computation of PI for a group of countries. Problems of construction of PI and remedial actions are addressed with the work of [4] as an example.

2 Construction of participation index

2.1 Framework on the concept of citizen participation

There is no agreement in operational definition of multidimensional participation index (PI). Active citizenship composite index (ACCI) considered participation in four dimensions viz. Political Life, Civil Society, Community Life and the Values and different number of sub-dimensions and indicators under each dimension [5]. Direct impact of participation in the plan-making process was found [6] in inclusion outcomes to reduce cultural barriers; mitigating poverty; increase economic opportunities and promote good governance, where degree of relationship varied across different indicators of social inclusion. In ref. [7] observed that theoretical framework of active ageing failed to identify the relevant contributing factors and barriers and need improved conceptual and analytical clarity. Moreover, empirical investigations of active ageing depend on health, education, good finances, etc. Thus, evaluation of active ageing for individuals may be an elusive goal for a large segment of older adults.

The European Charter on the Participation of Young People in Local and Regional Life (1992 and revised in 2003) is an international policy document approved by the Congress of Local and Regional Authorities of the Council of Europe (www.salto-youth.net/downloads/4-17-1510/Revised%20European%20Charter%20on%20the%20Participation%20of%20YP.pdf). The charter outlines the following 14 areas for involvement of young people:

1. Sport, leisure and associative life
2. Work and employment
3. Housing and transport
4. Education and training
5. Mobility and intercultural exchanges
6. Health
7. Equality for women and men
8. Young people in rural areas
9. Access to culture
10. Sustainable development and environment
11. Violence and crime
12. Anti-discrimination
13. Love and sexuality
14. Access to rights and law.

Similarly, the National Report on Youth Policy in Norway [8] summarizes a number of critical changes which impact on the lives of young people and has significant overlaps with the Council of Europe.

Socio-economic inequalities may restrict participation in the society and even undermines responsiveness and representativeness of people with disadvantaged backgrounds [9, 10]. Other factors influencing participation in positive fashions are increase in age, family income, education, etc. [11]. However, methodological issues like scaling, selection of weights and aggregation of indicators suffer from limitations for meaningful inter-country/inter-regional/inter-sample comparisons, classification, assessment of progress, etc. Statistical testing of equality of mean of participation across time and space can be done if assumptions of the techniques like normal distribution of the variables are satisfied. Effect of active participation to wellbeing and quality of life has been addressed by [12–14].

2.2 Data

To operationalize the concept, 63 basic indicators distributed over the following four dimensions and a number of sub-dimensions were considered. Details are given in Table 1.

It may be observed that indicators like *money donated* are related with per capita GDP, which is an indicator of Human Development Index (HDI). Similarly, *Women Participation in national parliament* overlaps with Global Gender Gap Index (GGGI); *Democracy* may coincide with trust, perception of inequality, which are components of Social Cohesion Index (SCI). Thus, PI is likely to have positive correlations with HDI, GGGI, and SCI.

2.3 Natures of data on selected indicators

– Membership to Political parties, Human rights organisations, Trade union organisations, and other organisations are in ratio scale usually given in percentages. However, addition of X% = $\frac{N_X}{D_X} \times 100$ and Y% = $\frac{N_Y}{D_Y} \times 100$ is not meaningful when $D_X \neq KD_Y$. For example, $\frac{8}{24}$ = 33.33% and $\frac{5}{10}$ = 50%. Combined percentage is $\frac{13}{34}$ = 38.23% and not average of 50% and 33.33% i.e. 41.67%

– Participation could be at different levels like very active, active, dormant, etc. Information about participation obtained by survey using 3-point or 5-point items are ordinal and needs appropriate method of aggregation since latent distance between "very active" and "active" is not same between "active" and "dormant". Thus, ordinal scores emerging from 3-point or 5-point items are not equidistant and hence, addition is not meaningful. $\overline{X} > \overline{Y}$ or $\overline{X} < \overline{Y}$ is meaningless for ordinal scales [15].

– Data on money donated (in ratio scale) may not be reliable.

– Voting Turnout in Parliament, Women Participation in national parliament are secondary data in percentages.

– Working in an organisation or association is a binary indicator of Yes–No type. Signing a petition, taking part in lawful demonstrations, boycotting products, contacted a politician, etc. say in last 12 months are in numbers. However, concepts of lawful demonstrations, ethical consumption, etc. vary and need to be defined properly without ambiguity.

– Non-organised help in the community may convey differently to different subjects. For better understanding of community organizing, [16] found it is useful to discuss what organizing is not. Functions of community organizers are different from community developers, social service professionals, lawyer, or other fields of community engagement.

– Indicators of dimension of Values need to be defined since they are subjective and sensitive too. Questioner for such indicators needs to be developed carefully avoiding leading items/questions.

– Correlation between a pair of indicators may vary. Very high correlation between two indicators gives rise to multi-colinearity and needs to be avoided.

2.4 Observations

The selected indicators may not be exhaustive since active citizenship is an evolving concept. IT-related interactions and participations need to be included. Increasing use of digital technology like internet, civic/political websites, etc. offering diverse forms of interactive participations, information sharing, etc. may reflect the relationship between citizen and state, which could also influence the social networks on citizens' willingness to participate in social governance. In ref. [17] used online and offline survey covering 15 provinces of China using questionnaires consisting of 1037 items including 773 valid items from online survey, and 250 items from offline survey. Here, the dependent variable "the citizens' willingness to participate in social governance was assessed by 4-points items like "not willing at all," "willing to participate in all the available activities," "willing to participate in the activities of easy access," and "willing to participate in the activities of easy access, with self-interests involved". Independent variables were social capital having three parts viz. social networks, social trust, and social norms and each was assessed by 4-point items like "how many neighbors do you greet frequently in your community" and "how many neighbors are considered to be your friends in your community" To asses "social trust," and "social norms," Likert scales were used generating Ordinal scores.

Scale to assess civic competence is multidimensional due to presence of different dimensions. Principal Component Analysis (PCA) or Factor Analysis (FA) of indicators or dimensions may result in several independent factors. Thus, finding total scores of a multidimensional scale may not be appropriate. For example, the manual of the 36-Item Short Form

Table 1 Dimensions and corresponding indicators

Dimensions	Sub-dimensions	Indicators
Political life	*Political parties*	Membership, participation, donating money, voluntary work, worked in political party/action group during last 12 months, donated money to political organisation/action group during last 12 months
		Voting turnout in parliament
		Women participation in national parliament
Civil society dimension	*Protest*	Working in an organisation or association, signing a petition, taking part in lawful demonstrations, boycotting products, ethical consumption, contacted a politician
	Human rights organisations	Membership, participation, donating money, voluntary work
	Trade union organisations	Membership, participation, donating money, voluntary work
	Environmental organisations	Membership, participation, donating money, voluntary work
Community life dimension	*Religious organisations*	Membership, participation, donating money, voluntary work
	Business organisations	Membership, participation, donating money, voluntary work
	Sports organisations	Membership, participation, donating money, voluntary work
	Cultural organisations	Membership, participation, donating money, voluntary work
	Social organisations	Membership, participation, donating money, voluntary work
		Non-organised help in the community
Dimension of values	*Human rights*	Immigrants should have same rights, law against discrimination in the work place, law against racial hatred
	Intercultural	Allow immigrants of different race group from majority, cultural life undetermined/enriched by immigrants, immigrants make country worse/better place
	Democracy	How important for a citizen to vote, How important for a citizen to obey laws, how important for a citizen to develop an independent opinion, how important for a citizen to be active in a voluntary org., how important for a citizen to be active in politics

Health Survey questionnaire (SF-36) provides no support to calculate total/overall score of SF-36 for an individual or a country due to several independent factors being measured by the scale (http://www.webcitation.org/6cfeefPkf)

2.5 Arithmetic aggregation

Addition of variables in ratio scale and variables in ordinal scale are problematic and difficult to interpret. Even if each of X and Y is in ratio scale, $X + Y = Z$ is meaningful if X and Y follow similar distribution (may be with different parameters) and probability density function (pdf) of Z can be derived as convolution of distributions of X and Y.

Summative scores of a scale as sum of item scores assign equal importance to the items and dimensions despite different item-total or dimension-total correlations, factor loadings for dimensions and items as observed from PCA. Summative scores also suffers from substitution effect (poor score in one dimension can be compensated by higher scores in other dimensions) and may mislead results [18, 19].

Better could be to transfer data on K-point items (K = 2, 3, 4, 5, ……..) to equidistant scores which can be further transformed to follow normal distribution with same range of item scores [20]. Here, sum of normally distributed item scores will also follow normal, facilitating estimation of the parameters from the data.

In addition, summative scores may result in a number of tied scores and reduce discriminating power of the scale. Quality of a scale in terms of psychometric properties like reliability, validity, discriminating value need to be reported.

2.6 Scaling

Raw scores of the indicators in different units and in different score ranges are usually scaled before aggregation. Min–Max function was used [4] where i-th indicator (X_i) was rescaled to unitless Y_i by $Y_i = \frac{X_i - Min(X_i)}{Max(X_i) - Min(X_i)}$ where $0 \leq Y_i \leq 1$

Such scaling indicates relative performance and not absolute performance of a country [21]. It depends significantly on X_{Max} and X_{Min} which may be unreliable outliers. If X is in percentage, $[Max(X_i) - Min(X_i)]$ may not be meaningful. Human Poverty Index (HPI) considers 3rd root and 4th root of average of figures in percentage for HPI-1 and HPI-2 respectively [22]. Min–Max transformation tends to overestimate the impact of indicators having small score ranges and changes distribution of the transformed scores and may have impact on the PI. Ranks of two countries may be influenced by performance of a third country [23]. Decrease in performances of the worst performing country may increase value of Y_i, even if X_i for the i-th country remains unchanged. If X_{Min} is changed, ranking and relative valuations may be changed due to change in marginal rates of substitution [24]. X–Y curve for a country is not linear since increase in Y per unit increase in X is different for different values of X. For example, consider an indicator X taking values: 58, 70, 80, 85, 90, 96. Clearly, $X_{Max} = 96, X_{Min} = 58$. Gain in Y due to increase of X from 80 to 90 is 0.2632. Similar gain from increasing X from 85 to 90 is 0.1032 which is less than 0.2632.

Better scaling ensuring unit less values is standardization by $Z = \frac{X - Mean(X)}{SD(X)} \sim N(0, 1)$ where $-\infty < Z_i < \infty$. Negative scores can be avoided by further transforming Z-scores to have a desired score range say 1 to 100. In PCA, original input variables are standardized to Z-scores.

Other illustrative transformations for scaling are:

- $Z_i = \frac{X_i}{\bar{X}} \times 100$. It is less robust to the influence of outliers and is linearly related to Proportionate Normalization where $Z_i = \frac{X_i}{\sum X_i}$

- $Z_i = \frac{X_i}{X_{Max}} \times 100$. Depends on X_{Max},

- For longitudinal data, $Y_i^t = \frac{X_i^t - X_i^{t-1}}{X_i^t} \times 100$ where t denotes time period or ratio with X_i^0, the base period as $\frac{X_i^t}{X_i^0} \times 100$ for each positive indicator/domain and $\frac{X_j^0}{X_j^t} \times 100$ where j denotes a negative indicator/domain.

- Logarithmic transformation of an indicator: $Y_i = ln(X_i)$ for $X_i \geq 0$. For the Income component, HDI (2010) used $Income_X = \frac{log_e^X - log_e^{(X_{Min})}}{log_e^{(X_{Max})} - log_e^{(X_{Min})}}$. Here, rate of increase of $Income_X$ is different for different values of X and $Income_X$ is not invariant under change of origin. Logarithmic transformation fails to satisfy desired properties like Translation Invariance and consistency in aggregation [25]. The index depends on the normalization methods applied to different

indicators. Moreover, non-linear logarithmic transformation may alter the structure of the data. $r_{Lifeexpectancy,HDI} > r_{Lifeexpectancy,GDP}$ [26]. But, the inequality got reversed on taking logarithmic transformations.

The scaling methods have advantages and disadvantages. There is no best method of scaling. Different normalization techniques resulted in different ranking lists [27]. Multi-criteria decision-making (MCDM) methods like Analytic Hierarchy Process (AHP), Data envelopment analysis (DEA), Benefit of the Doubt (BoD), etc. require no normalization.

2.7 Weighted sum

Weighted sum is a popular approach for aggregation where weights ($W_i\prime s$) are chosen to the indicators ($X_i\prime s$) and the composite index (CI) is obtained as $CI = \sum W_i X_i$ where $0 < W_i < 1$ and $\sum W_i = 1$ to satisfy the convex property. In ref. [4] used weighted sum to get dimension scores and further weights to the dimensions to compute PI.

There are different methods of selecting weights. Approaches to find weights to the indicators could be normative (subjective weights), data-driven (determined objectively) or hybrid [28]. Selected weights serve as 'trade-offs' and reflect relative importance. The amount of indicator 2 to be sacrificed to gain an extra unit of indicator 1 is reflected in the ratio $\frac{W_1}{W_2}$. Such trade-offs may not be meaningful when the indicators relate to economic growth (say GDP) and improvement in non-monetary areas like values against discrimination.

Weights of DEA are obtained satisfying relevant constraints of the Linear programming and deriving a single aggregate measure for each decision making unit (DMU). DEA results are sensitive to the choice of inputs and outputs. However, the best specification cannot be tested. Moreover, number of efficient units tends to increase with increase in number of inputs and output variables [29]. AHP finds weights of criteria based on $\frac{n(n-1)}{2}$ pair-wise comparisons for *n*-number of alternatives using 1 to 9 scales, which gets increased with increase in *n*. Inconsistency may take place since transitivity rule does not hold good for all elements of the matrix showing pair wise comparisons, ambiguity in judgement scale may not be avoided. AHP cannot be of much use to requirements priority [30].

Best–Worst Method (BWM) finds weights based on Tchebychev distance, requiring less pair-wise comparisons than AHP [31]. However, major issues of BWM are lack of threshold for the consistency ratio, ordinal consistency, and complex calculation process, especially for large *n*. There could be situations where there is no unique best and/or worst criterion/criteria (as required in BWM). Situations like ≥ 2 best or worst criteria cannot be solved easily by BWM. Using scale from 1 to 9 to determine the most important/preferred criterion (or the least important criterion) is subjective and is sensitive to the sample composition. BoD weights can be used for linear aggregation but not for non-linear/non-compensatory aggregation. The weight vector of AHP as factor loadings of the principal eigenvector has significant drawbacks [32]. For Economic Sentiment Indicator and environmental sustainability index, PCA and FA failed [33]. Weight vector $W = (W_1, W_2, \ldots., W_n)^T$ was proposed by [34] such that $\sum_{i=1}^{n} W_i = 1$ and variance of $Y = \sum_{i=1}^{n} W_i X_i$ is minimum. Instead of Xi's, if standardized scores $Z_{ij} = \frac{X_{ij} - \overline{X}_j}{S_{X_j}}$ are taken, the $r_{Y,Z_i} = r_{Y,Z_j} = \frac{1}{\sqrt{e^T R^{-1} e}}$ where $Y = \sum_{i=1}^{n} W_i X_i$ and R is the correlation matrix and $i \neq j$. In other words, Y is equi-correlated with $Z_i\prime s$.

However, different methods of selection of weights may affect PI differently and may not ensure rank robustness. CI as weighted sum do not discuss about variance of the weighted sum and correlation of CI with the chosen indicators. Equal weights were applied to the dimensions and indicators within each sub-dimension to compute Active Citizenship Composite Indicator [4]. Equal weighting with perfect substitutability may reverse country ranks [35]. No weight or equal weights are wrong and no weighting system is above criticism [36]. Just like no consensus of the perfect weighting scheme, there is no ideal aggregation scheme [37]. Thus, it is desirable to construct PI without considering weighted sum.

3 Proposed methods

3.1 Arithmetic aggregation (Method-1)

Let X_{ij} be the raw ordinal score of a country in the *i*-th item for choosing the *j*-th response-category where *higher score ⇔ higher participation*. For a 5-point item, weighted score (WS) $= \sum \sum W_{ij} X_{ij}$ where $W_{ij}\prime s$ are different for different levels of the item satisfying $W_{ij} > 0$ and $\sum_{j=1}^{5} W_{ij} = 1$.

Scores of the *i*-th item will be equidistant and monotonic if $W_{i1}, 2W_{i2}, 3W_{i3}, 4W_{i4}$ and $5W_{i5}$ forms an arithmetic progression with positive common difference (α).

For the i-th item, find maximum (f_{imax}) and minimum frequency (f_{imin}) of the levels. Find initial weights $\omega_{ij} = \frac{f_{ij}}{n}$.

Arrange the ω_{ij}'s so that $\omega_{i1} < \omega_{i2} < \omega_{i3} < \omega_{i4} < \omega_{i5}$ where $\omega_{i1} = \frac{f_{imin}}{n}$ and $\omega_{i5} = \frac{f_{imax}}{n}$.

Let intermediate weight $W_{i1} = \omega_{i1}$

Compute α as $\alpha = \frac{5f_{imax} - f_{imin}}{4n}$ since $W_{i1} + 4\alpha = 5W_{i5}$

Other intermediate weights are $W_{i2} = \frac{\omega_{i1} + \alpha}{2}$, $W_{i3} = \frac{\omega_{i1} + 2\alpha}{3}$; $W_{i4} = \frac{\omega_{i1} + 3\alpha}{4}$; and

$W_{i5} = \frac{\omega_{i1} + 4\alpha}{5}$. Get final weights $W_{ij(Final)} = \frac{W_{ij}}{\sum_{j=1}^{5} W_j}$ enabling $\sum W_{ij(Final)} = 1$ and

$j.W_{j(Final)} - (j-1).W_{(j-1)(Final)} = $ constant (different for different items).

3.1.1 Observations

i) $W_{j(Final)}$ are based on empirical probabilities.
ii) $f_{ij} = 0$ for a particular level of an item, can be taken as zero value for scoring Likert items as weighted sum.
iii) Generated scores (E) as weighted sum are continuous, equidistant and monotonic in ratio scale.
iv)   Method-1 is applicable for items with different number of response-categories including binary items.

Standardize E-scores of the i-th item by $Z_{ij} = \frac{X_{ij} - \overline{X}_i}{SD(X_i)} \sim N(0, 1)$.

Take linear transformation of Z_{ij} to P-scores by:

$$P = (99) * \left[\frac{(Z_{ij} - Min(Z_{ij}))}{Max(Z_{ij}) - Min(Z_{ij})} \right] + 1 \tag{1}$$

For the i-th item, $P_i \sim N(\mu_i, \sigma_i^2)$ and $1 \leq P_i \leq 100$ where estimates of μ_i and σ_i^2 are obtained from the data. P-score of an item as per Eq. (1) can be obtained irrespective of length of scale and width of items.

For the indicators in ratio scales, raw scores can be standardized and use (1) to follow normal.

Domain score of a country is taken as sum of normally distributed P-score of relevant items which will follow normal with mean $\sum_i \mu_i$ and SD $= \sqrt{\sum \sigma_i^2 + 2\sum_{i \neq j} Cov(P_i, P_j)}$. Similarly, scale score is sum of domain scores also following normal.

3.1.2 Properties

1. Domain scores (D_i) and scale scores (S_i) of the i-th country are continuous, monotonically increasing, each following normal distribution. Normality ensures meaningful computation of arithmetic average, SD, correlation, etc. and facilitates statistical analysis under parametric set up including unbiased estimates of population mean (μ), population variance (σ^2), confidence interval of μ, and testing of null hypothesis $H_0 : \mu_1 = \mu_2$ or $H_0 : \sigma_1^2 = \sigma_2^2$ etc. across time and space.

2. $Var(S_i) > \sum Var(D_i) \Rightarrow$ Positive correlations for most pair of dimensions.

3. Progress of the i-th country in successive time-periods can be quantified by $\frac{S_{i(t)} - S_{i(t-1)}}{S_{i(t-1)}} \times 100$, which also quantifies responsiveness of the scale and effectiveness of a implement policy measure. Deterioration is indicated if $\frac{S_{i(t)} - S_{i(t-1)}}{S_{i(t-1)}} \times 100 < 0$

Similarly, progress for a group of countries is reflected if $\overline{S_{i(t)}} > \overline{S_{i(t-1)}}$. Normally distributed S_i helps to test $H_0 : \mu_{S_t} = \mu_{S_{(t-1)}}$ and H_0: $Progress_{(t+1)overt} = 0$. Deterioration if any may be probed to find extent of deterioration in concerned domain(s) for possible corrective actions.

4. The plot of progress/deterioration of a country at various time points can be used to compare PI from the base period.

5. Policy makers are interested to know elasticity of each dimension as change in S-score per unit change in a dimension score i.e. $\frac{\nabla S}{\nabla D_i}$. The dimensions can be ranked in terms of $\frac{\nabla S}{\nabla D_i}$

6. Normally distributed scores satisfy the assumptions of PCA and enable to find factorial validity in terms of ratio of the first eigenvalue to the sum of all eigenvalues i.e. Factorial Validity $= \frac{\lambda_1}{\sum \lambda_i}$, where λ_1 is the highest eigenvalue reflecting the main factor for which the scale was developed[38] and accounts for $\frac{\lambda_1}{\sum \lambda_i} \times 100$ percent of overall variability. Such

factorial validity avoids the problems of construct validity and selection of criterion scale. Tracy–Widom (TW) statistic can be used to test significance of the largest or other eigenvalues [39].

7. Normality helps to estimate variance of subclass, scale and each item and estimated Cronbach alpha for a domain/subclass at population level as

$$\hat{\alpha} = (\frac{n}{n-1})(1 - \frac{Sum\ of\ estimates\ of\ variance\ of\ items\ in\ the\ sub-class)}{Estimate\ of\ variance\ of\ the\ sub-class}) \tag{2}$$

Reliability of scale (r_{tt}) consisting of K-number of dimensions can be obtained as a function of dimension reliabilities by

$$r_{tt} = \frac{\sum_{i=1}^{K} r_{tt(i)}S_{Xi} + \sum_{i=1,i\neq j}^{K} \sum_{j=1}^{K} 2COV(X_i, X_j)}{\sum_{i=1}^{K} S_{Xi} + \sum_{i=1,i\neq j}^{K} \sum_{j=1}^{K} 2COV(X_i, X_j)} \tag{3}$$

where $r_{tt(i)}$ and S_{xi} denote respectively reliability and SD of the i-th dimension [34].

8. Discriminating value of a scale indicates ability of the scale to distinguish between countries that have different degrees of the underlying construct. Discriminating value of an item ($Disc_i$) and test ($Disc_{Test}$) can be computed by Coefficient of variation (CV) where $Disc_i = \frac{SD_i}{mean_i}$ and $Disc_{Test} = \frac{SD_{Test}}{Mean_{Test}}$. Cronbach α and $Disc_{Test}$ (with m-items) are related by

$$\alpha = (\frac{m}{m-1})(1 - \frac{\sum_{i=1}^{m} \overline{X}_i^2 \cdot Disc_i^2}{\overline{X}^2 \cdot Disc_T^2}) \tag{4}$$

Since, variance of the i-th item $S_{X_i}^2 = \overline{X}_i^2 \cdot Disc_i^2 \,\forall\, i = 1, 2, \ldots, m \Rightarrow \sum_{i=1}^{m} S_{X_i}^2 = \sum_{i=1}^{m} \overline{X}_i^2 \cdot Disc_i^2$ and Test variance $S_X^2 = \overline{X}^2 \cdot Disc_T^2$

It can be proved that

$$(Disc_{Test})^2 = \frac{CV_{Truescores}^2}{r_{tt}} \text{ where } r_{tt} = \frac{S_T^2}{S_X^2} \tag{5}$$

Thus, test reliability and $Disc_{Test}$ are related by a negative non-linear relationship.

9. Quartile clustering helps in classification of a group of countries in four mutually-exclusive classes Q_1, Q_2, Q_3, Q_4. Quartile clustering of scale scores following normal distribution may be adopted because it is simple, appealing, provides well-defined cut-off scores for the four mutually exclusive classes and assigns equal probability to each quartile/class

$$\int_0^{Q_1} f(x)dx = \int_{Q_1}^{Q_2} f(x)dx = \int_{Q_2}^{Q_3} f(x)dx = \int_{Q_3}^{Q_4} f(x)dx \tag{6}$$

Similar approach may be used for classification to five classes (pentiles), ten classes (deciles), etc.

For normally distributed dimension scores, a given score X_0 in dimension-1 will be equivalent to a score of Y_0 in dimension-2 if

$$\int_{-\infty}^{X_0} f(x)dx = \int_{-\infty}^{Y_0} g(y)dy \tag{7}$$

where $f(X)$ and $g(Y)$ denote normal probability density function (pdf) of transformed scores of dimension-1 and dimension-2 respectively. The Eq. (7) can be solved using Standard Normal probability table. It helps to find all combinations of $\{X_0, Y_0\}$ including cut-off scores of two scales.

3.2 Geometric aggregation (Method-2)

Let X_{i0} denote value of the indicator in the base period. Unit free ratio $\frac{X_{it}}{X_{i0}}$ indicates progress or decline of the country with respect to the *i*-th indicator at *t*-th time period over the base period. PI for the *c*-th time period (current period) is defined as the Geometric mean of *n*-indicators i.e.

$$PI_{c0} = \sqrt[n]{\frac{X_{1c}, X_{2c}, \ldots \ldots .., X_{nc}}{X_{10} X_{20} \ldots \ldots .. X_{n0}}} \times 100 \tag{8}$$

or equivalently by avoiding the *n*-th root,

$$PI_{c0} = \frac{X_{1c}, X_{2c}, \ldots \ldots .., X_{nc}}{X_{10} X_{20} \ldots \ldots .. X_{n0}} \times 100 \tag{9}$$

Equation (9) can also be written in terms of dimensions scores as

$$DI_{c0} = \frac{D_{1c} D_{2c} \ldots \ldots . D_{kc}}{D_{10} D_{20} \ldots \ldots .. D_{k0}} \times 100 \tag{10}$$

$$\text{Clearly,} \, logPI_{c0} = \frac{1}{n} \sum_{i=1}^{n} log \frac{X_{ic}}{X_{i0}} \tag{11}$$

Here, $\frac{X_{it}}{X_{i0}} > 1$ indicates progress and $\frac{X_{it}}{X_{i0}} < 1$ indicates decline assuming *higher score ⇔ higher participation*. The indicators for which $\frac{X_{it}}{X_{i0}} < 1$ are critical and can be identified as the indicators for which $\frac{X_{it}}{X_{i(t-1)}} < 1$ for deciding appropriate action plan. $\frac{PI_{i_t}}{PI_{i_{(t-1)}}} > 1$ implies overall improvement made by the *i*-th country in period *t* over (*t*-1)-th period.

Note that each of (9) and (10) can be applied for all types of indicators in ratio scales or in ordinal scales and considers all chosen indicators including those in percentages and depicts overall improvement/decline in the current year over the base year by a continuous function which is symmetric over its arguments and increases monotonically. Replacing the base period vector by the vector for the previous year will give improvement of PI on year-to-year basis.

3.2.1 Properties

Each of (9) or (10) satisfy the following desired properties:

1. Independent of order of the chosen indicators or dimensions and independent of change of scale
2. Increase of say 1% in $\frac{X_{it}}{X_{i0}} \Rightarrow 1\%$ increase in the PI if all others are unchanged. Thus, the curve showing gain in $\frac{X_{it}}{X_{i0}}$ and gain in PI is linear.
3. Avoids scaling, selection of weights and reduces level of substitutability between component indicators.
4. Not affected much by outliers and thus produce no bias for developed or under-developed countries.
5. Satisfaction of time-reversal test and formation of chain indices by Method-2 may help inter-country comparison over time by tracking the path of overall progress registered by a country.
6. Facilitates construction of separate indices for each domain by focusing on indicators related to that domain without further weights for domains.
7. Easy to find relative importance of each indicator. The ratios for which $\frac{X_{it}}{X_{i0}} < 1$ are the critical areas requiring attention.
8. Reliability, factorial validity, discriminating value of *PI* or a dimension can be found taking logarithm on both sides of Eq. (10).
9. Facilitates estimation of population *GM*, standard error of the *GM* and confidence interval of *GM* [40] since $logGM = \frac{1}{n} \sum_{i=1}^{n} logY_i$ where $Y_i = \frac{X_{it}}{X_{i0}}$. Geometric standard deviation (GSD), is given by log $S_{GM} = [\frac{1}{n} \sum_{i=1}^{n} (logY_i - logGM)^2]^{\frac{1}{2}}$ where S_{GM} denotes GSD. This implies log (GSD of $Y_1, Y_2, \ldots \ldots, Y_n$) = usual SD of $logY_1, logY_2, \ldots \ldots, log Y_n$

Population estimate of *GM* is sample *GM* for large sample and estimate of standard error of the *GM* is $SampleGM.(\frac{\log S_{GM}}{\sqrt{n-1}})$.
Upper and lower limit of the confidence interval of *GM* for (1−α) % are (e^{U}, e^{L}) where $U = logGM + S_{m}.t_{(\frac{\alpha}{2},df)}$ and $L = logGM - S_{m}.t_{(\frac{\alpha}{2},df)}$

Thus, significance tests of hypotheses regarding equality of *GM's* can be performed across time and space using conventional *t*-tests on the logarithms of the observations.

4 Benefits of Method-1 and Method-2

The proposed *PI* can be constructed even for skewed longitudinal data over long time periods and also for two time periods. Thus, it helps in meaningful comparison of a set of regions/countries with reduced level of substitutability among the indicators/dimensions. For comparison of sub-groups of a country say ethnic groups, religious groups, disadvantage groups, elderly people, etc., the proposed index can also be applied with pre-determined indicators. Method-2 offers a simple solution to assess PI even for a single country from the base period without resorting to group data. Both methods can be well used for undertaking Cluster analysis and classification of countries. The graph of progress/decline of PI over time, for a country will help to assess impact of various socio-economic measures across time.

In addition to citizen participation, multidimensional CI as proposed in this article can also be applied to gender equality, quality of life, and socioeconomic well-being [41, 42]. It would be useful to comment on other composite indicators used in this type of studies, such as the synthetic distance indicator DP (2) [43] or the two methods proposed here to estimate gender equality. Empirical comparison of the two methods could be made and the most appropriate one selected for the estimation of citizen participation.

5 Limitations

1. Introduction of new indicator requires estimation of value of that indicator in the base period and subsequent periods.
2. The method fails if an indicator is ≤ 0.

6 Discussion

Focusing on methodology of assessing PI, the paper describes two assumption free methods avoiding scaling. Method-2 also avoids selection of weights. Thus, the Method-2 avoids the problems of rank robustness specific to selection of weights [35]. Normal distributions of two or more groups by Method-1 are likely to give rise to lower value of Ginis coefficient indicating equality. Conversion of discrete and ordinal scores of *k*-point items (X_i) to continuous, monotonically increasing equidistant scores (E_i) by weighted sum where weights to different response-categories of different items are in ratio scales. Linear transformations of *E*-scores to *Z*-scores and *P*-scores can be added with indicators measured in ratio scales.

Benefits of both the method include:

– Meaningful arithmetic aggregation (Method-1). Multiplicative aggregation in Method-2 can be converted to additive model by taking logarithms.
– Normally distributed dimension scores (D_i) and *PI* satisfy many desired properties like meaningful comparisons, ranking and classification of set of countries, estimation of population parameters and testing of statistical hypothesis. However, normality of *PI* to be tested for Method-2 by say Anderson –Darling test.
– Identify critical dimensions showing deterioration with time and thus, draw attention of the policy makers for initiation of corrective actions
– Find effect of small change in *i*-th dimension to *PI* by elasticity as ratio of change in *PI* due to unit change in a dimension. Such elasticity can be used to rank the dimensions.
– Possible to assess Progress/decline of *PI* at successive time periods for monitoring of effect of policies and strategies. Statistical hypothesis of significance of progress/decline of *PI* at two different time periods can be undertaken since ratio of two normally distributed variable follows χ^{2} distribution.
– High reliability of the scale (r_{tt}) indicates rank robustness, which can be quantified by Spearman's rank correlation (ρ).

Thus, the proposed measures satisfying the desired properties is an improvement over the existing methods

7 Conclusion

The paper contributes to improve scoring of PI avoiding major limitations of ordinal scores and facilitating analysis under parametric set up for meaningful comparisons. Policy makers and researchers can take advantages of the proposed methods of arithmetic aggregation without weights or multiplicative aggregation without scaling and choosing weights. Both methods satisfy desired properties. Method-2 offers more generalized approach satisfying time reversal test and formation of chain indices. However, test of normality is required for this method unlike the Method-1 which ensures normally distributed scores. The proposed methodologies are innovative and contribute to estimation of citizen participation in their respective territories or countries. Empirical studies may be undertaken to find correlations of PIs by each proposed method and generalization of findings.

Declarations

References

1. Hoskins B (2006) Draft framework on indicators for active citizenship. Ispra: CRELL, 31.
2. Yani AA, Hidayat AR. What is the citizenship quality of our community? Measuring active citizenship. Public Adm. 2018. https://doi.org/10.17323/1999-5431-2018-0-6-119-133.
3. OECD. Handbook on constructing composite indicators: methodology and user guide. Paris: OECD Publishing; 2008.
4. Hoskins B, Mascherini M. Measuring active citizenship through the development of a composite indicator. Soc Indic Res. 2009;90(3):459–88. https://doi.org/10.1007/s11205-008-9271-2.
5. Sonia L, Bober M, Helsper EJ. Active participation or just more information? Inf Commun Soc. 2005;8(3):287–314.
6. Nwachi L. Relationship between participation and social inclusion. Open J Soc Sci. 2021;9:46–77. https://doi.org/10.4236/jss.2021.910004.
7. Jensen PH, Skjøtt-Larsen J. Theoretical challenges and social inequalities in active ageing. Int J Environ Res Public Health. 2021;18(17):9156. https://doi.org/10.3390/ijerph18179156.
8. Ministry of Children and Family affairs, Norway. National Report on Youth Policy in Norway, Report by the International Team of Experts Presented to the Joint Council on Youth. (CMJ) Strasbourg. 2004.
9. Donbavand S, Hoskins B. Citizenship education for political engagement: a systematic review of controlled trials. Soc Sci. 2021;10:151.
10. Hoskins B, Huang L, Arensmeier C. Socioeconomic inequalities in civic learning in Nordic schools: identifying the potential of in-school civic participation for disadvantaged students. In: Biseth H, Hoskins B, Huang L, editors. Northern lights on civic and citizenship education. IEA research for education, vol. 11. Cham: Springer; 2021. https://doi.org/10.1007/978-3-030-66788-7_5.

11. Lauglo J, Oia T. Education and civic engagement among Norwegian youth. Policy Future Educ. 2008;6(2):203–23. https://doi.org/10.1080/13691180500259103.

12. Constantinescu M, Orîndaru A, Căescu ȘC, Pachiṭanu A. Sustainable development of urban green areas for quality of life improvement—argument for increased citizen participation. Sustainability. 2019;11(18):4868.

13. van Dinter M, Kools M, Dane G, Weijs-Perrée M, Chamilothori K, van Leeuwen E, van den Berg P. Urban green parks for long-term subjective well-being: empirical relationships between personal characteristics, park characteristics, park use, sense of place, and satisfaction with life in The Netherlands. Sustainability. 2022;14(9):4911.

14. Reyes-Riveros R, Altamirano A, De La Barrera F, Rozas-Vásquez D, Vieli L, Meli P. Linking public urban green spaces and human well-being: a systematic review. Urban For Urban Green. 2021;61:127105.

15. Hand DJ. Statistics and the theory of measurement. J R Statist Soc A. 1996;159(3):445–92.

16. Schutz A, Sandy MG. What isn't community organizing? In: Schutz A, Sandy MG, editors. Collective action for social change. New York: Palgrave Macmillan; 2011. https://doi.org/10.1057/9780230118539_3.

17. Nan R, Ouyang F. Influence of social networks on citizens' willingness to participate in social governance: evidence from China. Complexity. 2020;2020(3819514):16. https://doi.org/10.1155/2020/3819514.

18. Ray AK. Measurement of social development: an international comparison. Soc Indic Res. 2008;86(1):1–46.

19. Mikulić J, Kožić I, Krešić D. Weighting indicators of tourism sustainability: a critical note. Ecol Ind. 2015;48:312–4.

20. Chakrabartty SN. Equidistant likert as weighted sum of response categories. Cult Edu Soc. 2023;14(1):75–92. https://doi.org/10.17981/cultedusoc.14.1.2023.04.

21. Sava AM. Weighting method for developing composite indices: application for measuring sectoral specialization. J Appl Quant Methods. 2016;II(3):77–84.

22. UNDP. Human development report. New York: Oxford University Press; 2007.

23. Kasparian J, Rolland A. OECD's Better life index: can any country be well ranked? J Appl Stat. 2012;39(10):1–8. https://doi.org/10.1080/02664763.2012.706265.

24. Seth S, Villar A. Measuring human development and human deprivations. OPHI Working Paper 110, University of Oxford. 2017.

25. Chakrabartty SN. Wellbeing index avoiding scaling and weights. Cult Edu Soc. 2021;12(2):181–204. https://doi.org/10.17981/cultedusoc.12.2.2021.11.

26. Kovacevic M. Review of HDI Critiques and Potential Improvements, The Human Development Research Paper 2010/33, UNDP. 2011.

27. Mhlanga ST, Lall M. Influence of normalization techniques on multi-criteria decision-making methods. J Phys. 2022. https://doi.org/10.1088/1742-6596/2224/1/012076.

28. Decancq K, Lugo MA. Weights in multidimensional indices of wellbeing: an overview. Economet Rev. 2009;32(1):7–34.

29. Berg S. Water utility benchmarking: measurement, methodology, and performance incentives, International Water Association. London: IWA Publishing; 2010.

30. Shanmuganathan M, Karthikeyan R. Limitations of AHP as applied to requirements prioritization. Artif Intell Syst Mach Learn. 2016;8(4):137–40.

31. Rezaei J. Piecewise linear value functions for multi-criteria decision-making. Expert Syst Appl. 2018;98:43–56.

32. Li C, Sun Y, Jia Y, Li H. An improved ranking approach to AHP alternatives based on variable weights. IEEEXplore. 2008. https://doi.org/10.1109/WCICA.2008.4594220.

33. Nardo M, Saisana M, Saltelli A, Tarantola S, Hoffmann A, Giovannini E. Handbook on Constructing Composite Indicators: Methodology and User Guide. Paris (France). 2008.

34. Chakrabartty SN. Better use of scales as measuring instruments in mental disorders. J Neurol Res Rev Rep. 2020;2(3):1–7. https://doi.org/10.47363/JNRRR/2020-(2)128.

35. Permanyer I. Assessing the robustness of composite indices rankings. Rev Income Wealth. 2011;57(2):306–26. https://doi.org/10.1111/j.1475-4991.2011.00442.x.

36. Greco S, Ishizaka A, Tasiou M, Torrisi G. On the methodological framework of composite indices: a review of the issues of weighting aggregation, and robustness. Soc Indic Res. 2019;141:61–94. https://doi.org/10.1007/s11205-017-1832-9.

37. Arrow KJ, Raynaud H. Social choice and multicriterion decision-making. Cambridge: MIT Press; 1986.

38. Parkerson HA, Noel M, Gabrielle MP, Samantha F, Katz J, Asmundson GJG. Factorial validity of the English-language version of the pain catastrophizing scale-child version. J Pain. 2013;14(11):1383–9. https://doi.org/10.1016/j.jpain.2013.06.004.

39. Nadler B. On the distribution of the ratio of the largest eigenvalue to the trace of a Wishart matrix. J Multivar Anal. 2011;102:363–71.

40. Alf EF, Grossberg JM. The geometric mean: confidence limits and significance tests. Percept Psychophys. 1979;26(5):419–21.

41. Chakrabartty SN. Methodological issues: gender related indices. Discov Glob Soc. 2023. https://doi.org/10.1007/s44282-023-00012-w.

42. Chakrabartty SN. Quality of life for patients with heart disease: limitations and remedial actions. Indian J Pharm Pract. 2023;16(3):232–7. https://doi.org/10.5530/ijopp.16.3.38.

43. Arechavala NS, Trapero BP. Synthetic indicators of the quality of life in Europe. In: Michalos AC, editor. Encyclopedia of quality of life and well-being research. Dordrecht: Springer; 2014. https://doi.org/10.1007/978-94-007-0753-5_3729.