Construction of a Reflective Quality of Life Index for Gauteng Province in South Africa

Samy Katumba[1] · **Julia de Kadt[1]** · **Mark Orkin[2]** · **Paul Fatti[3]**

Abstract

This paper demonstrates a simple methodological sequence for the statistical calculation of a context-sensitive quality of life index especially suitable for use in the Global South, i.e. lower and middle income countries. We draw on the large (n = 24,889), area-sampled, survey dataset of the fifth Quality of Life Survey (QoL V, 2017/18) of South Africa's Gauteng province, conducted by the Gauteng City-Region Observatory. Using overtly exploratory analysis, and a fully reflective conception of indicators, we apply exploratory factor analysis (EFA) in two stages to generate our model, while defending our approach methodologically and empirically and indicating its philosophical foundations. We contrast this approach to previous analyses using the QoL I (2010) data, which defined their dimensions and assign their indicators in advance, drawing on literature and indexes largely from the Global North. We ran series of EFAs on 60 longitudinally available variables from the QoL V data set. This allowed us to determine the optimal number of factors/dimensions, guided by established criteria and the interpretability of the grouped indicators. Each dimension score is calculated arithmetically, using the indicators' factor loadings as weights. Then the dimensions, weighted by their eigenvalues, are aggregated into the composite index, scaled to run from 0 to 100. The resulting seven-dimension, 33-indicator model was validated through confirmatory factor analysis (CFA) on the QoL V data; and its configural invariance, i.e. whether the pattern of indicators amongst the factors holds over time, was checked against the two previous QoL datasets and confirmed. This validates the ability of the approach outlined to generate a stable index. Analysis of the resulting quality of life index by race and municipality reveals trends consistent with the South African context. Overall, the White population group has the highest measured quality of life index, followed in turn by the Indian/Asian, Coloured and Black African population groups. Changes in the quality-of-life ranking of the nine municipalities comprising Gauteng province may be observed over time, using the validated previous models. The value of the exploratory approach in enabling context to influence the index construction is highlighted by the emergence of a distinct dimension of bottom-up political voice, which is relevant for democratic governance.

✉ Samy Katumba
   samykabangu@gmail.com

Extended author information available on the last page of the article

1 Introduction

Enhancing quality of life is a key objective of government and civil society globally. The creation of a composite quality of life index provides a statistically founded tool to measure progress and challenges in quality of life, creating a set of scores which is relatively straightforward to understand and interpret (Mazziotta & Pareto, 2017). To date, however, much of the work on creation of these indexes is driven by theoretical and empirical work shaped in the Global North. We argue that the creation of an index outside of this context must be empirically informed by an understanding of the relevant particularities of the context in question, rather than by application of an externally derived theoretical framework.

In this paper, we motivate the methodological and philosophical appropriateness of our approach, prior to using large-sample survey data from Gauteng province, South Africa to demonstrate a relatively straightforward statistical sequence for the calculation of a context-sensitive quality of life index. A large-scale, area-sampled Quality of Life Survey (QoL) has been conducted biennially since 2009 by the Gauteng City-Region Observatory (GCRO). To generate the index model, we apply exploratory factor analyses (EFA) in two stages to 60 longitudinally available variables in the data from this survey's fifth iteration, QoL V (2017/18) n = 24,889. This yields a seven-dimensional, 33-indicator model, which is validated using confirmatory factor analysis (CFA) on the QoL V dataset. Configural invariance is additionally confirmed using the two previous survey datasets (i.e. QoL IV (2015/16) and QoL III (2013)), suggesting that the index model can be used over time to generate longitudinally comparable results.

Following identification and validation of the model, each dimension score is calculated arithmetically, using the indicators' factor loadings as weights. These dimension scores, weighted by their eigenvalues, are then aggregated into a composite index, scaled to run from 0 to 100. We compare this construction of the index to the GCRO's previous approach to index calculation (Everatt, 2017), and identify noteworthy differences in both the number and constitution of dimensions. We also highlight the importance of the exploratory approach in generating a context-sensitive grouping of indicators, for example enabling the emergence of a distinct dimension representing bottom-up political voice. While not a standard feature of many quality of life indexes, its importance for democracy has been emphasised by Stiglitz et al. (2009).

The aim of this study is therefore to devise a data-driven methodology for constructing a quality of life index in the context of South Africa (and Gauteng province in particular), a middle income country. But we expect that the straightforward methodological approach we outline to data-handling, model generation, validation, and index calculation, is suitable for application elsewhere in other low and middle income countries.

2 The GCRO's Biennial QoL Surveys

### 2.1 Importance of the Surveys in the South African Context

Contemporary South Africa continues to struggle with many objective legacies bequeathed by the apartheid era (Statistics South Africa, 2019): racially aligned poverty, inequality, unemployment and spatial segregation. These come with behavioural and subjective concomitants including high levels of crime and violence (South African Police Service, 2017), widespread ongoing protests and pervasive mental distress (Mungai & Bayat, 2019). These challenges have proved intractable, and improvement has been gradual and uneven. These circumstances have been the subject of plans at national level (National Planning Commission, 2012) and below, along with innumerable local and international academic and stakeholder studies. The vast range of resulting official and academic statistics that are continually reported in the media make it nigh impossible for political principals, civil society, and laypersons alike to achieve an informed overall view of progress and change.

In this context, quality of life studies have played and continue to play an essential role in selecting salient indicators, integrating them into suitable dimensions and domains, and supplying a one-number index that can be tracked and differentiated, regionally and across relevant social groupings. Notable quality of life surveys were conducted and repeated in South Africa by independent academics during and since the apartheid years (Møller, 2014), and have been sustained in biennial national studies since the onset of democracy by the government-funded Human Sciences Research Council (Rule, 2007), while numerous once-off quality of life studies with particular topical or geographical foci continue to be mounted (Møller & Roberts, 2014).

Encouraged by the Stiglitz et al. report (2009), which both emphasised the importance of quality of life in sustainable development and stressed the feasibility of sample surveys of objective and subjective aspects of quality of life, the GCRO was established as a partnership of the University of the Witwatersrand, Johannesburg, the University of Johannesburg, the Gauteng Provincial Government and organised local government. A key objective of the GCRO is to provide government, researchers and civil society with regular and academically sound research and insight into key contributors to individual and societal wellbeing, including household circumstances, service delivery, perception of government, social attitudes, health, life satisfaction etc. This information has been collected by GCRO since 2009 through five large-scale, biennial, carefully area-sampled, surveys of randomly selected adult respondents in households across the province (Everatt, 2017; GCRO, 2014, 2016, 2019; de Kadt et al., 2019). To provide an accessible tool for rapid assessment of overall wellbeing in the province, survey data was used to calculate a single-number QoL index score from the outset.

Because of the variety of official and academic user needs, these surveys have had to be 'omnibus' in nature. Nevertheless they are named the Quality of Life Surveys, from QoL I in 2009, to QoL V in 2017/18. (The results of QoL VI were released in 2021, following completion of the analysis presented here.) The surveys carry both core and rotating modules that include, but are not confined to, variables relevant to the typical dimensions of quality of life identified by Stiglitz et al. (2009). The original range of questionnaire items, including those devoted or applicable to quality of life, was informed by an overview of relevant quality of life and other surveys (Jennings, 2012).

Given the breadth of coverage, the questionnaires are long, taking approximately an hour each on average (Orkin, 2020). So they have to be conducted face-to-face. Gauteng is
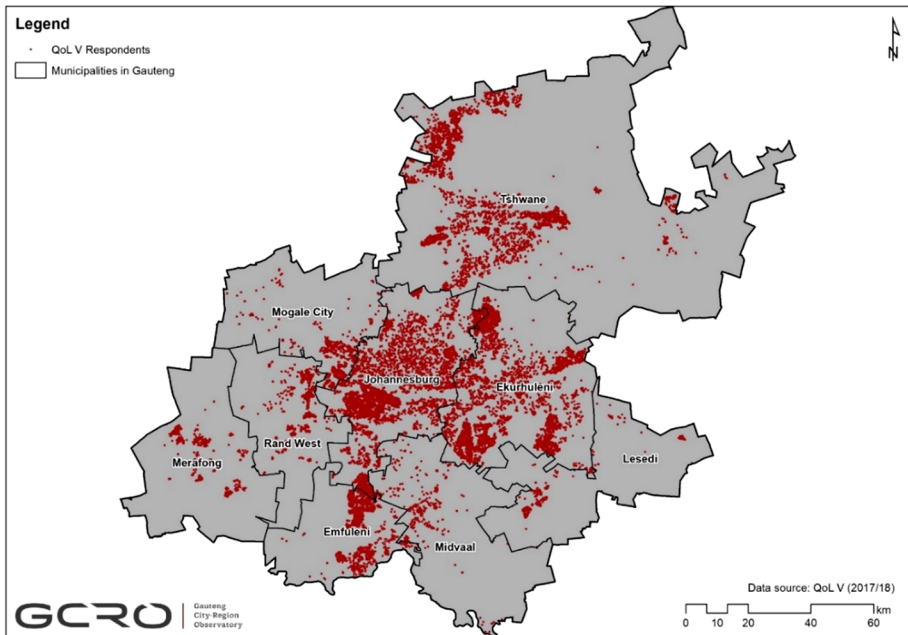
Fig. 1 GCRO QoL Survey V (2017/18): Spatial distribution of respondents (Map by Christian Hamann)

South Africa's most densely populated province, with its 15 million residents comprising a quarter of the country's population (Statistics South Africa, 2020), yet contained in only 1.4% of the country's land area. Even so, its residents live in dramatically different milieux, from urban houses and apartments through to effectively racially segregated 'townships'[1] enduring from the apartheid era, dense informal settlements, and even some still-rural areas (Statistics South Africa, 2019). As a result, the surveys are expensive to mount, and achieving specified coverage and response rates is arduous. Access to sampled households sometimes requires multiple repeat visits out of working hours. The spatial sample distribution for the GCRO QoL Survey V (2017/18), stratified across the nine municipalities of Gauteng, and drawing from the 529 wards, is mapped in Fig. 1.

2.2 The Original Approach to Generating the Quality of Life Index

Over time, the GCRO QoL surveys have steadily expanded in scope and sample: from 187 questions and 6,600 households in QoL I to 248 questions and 24,800 households in QoL V (Orkin, 2020). A core set of 123 questions has been sustained across survey iterations. To create the original QoL index, guided by a review of the literature, a generous subset of 54 questions arranged into ten domains was specified at the outset (Everatt, 2017). By QoL V the number of questions feeding into the QoL index had increased slightly to 58.

---

[1] In South Africa, 'townships' refer to racially segregated, densely populated areas, typically on the out-skirts of cities, in which Black African, Coloured or Indian population groups were constrained to live during the apartheid period. Examples around Johannesburg are Soweto for Black Africans, Eldorado Park for Coloureds, and Lenasia for Indians. They remain largely racially homogeneous.

Left model (10-Factor QoL Index)

10-Factor QoL Index	QoL index code	Variable
Infrastructure Dimension	i1	Has been an improvement in the community
	i2	Water is usually/always clean
	i3	Flush toilet facility
	i4	Piped water source
	i5	Electricity supply
	i6	Rubbish disposal
	i7	Haven't had water/electricity cut off
Socio-political Attitudes Dimension	p1	Disagree that politics is a waster of time
	p2	Agree that elections were/will be free and fair
	p3	Agree judiciary is free from government influence
	p4	other
	p5	Believe foreigners should be allowed to stay
	p6	Satisfaction: National government
	p7	Satisfaction: Provincial government
	p8	Satisfaction: Local municipality
	p9	Most government officials are doing their best?
	p10	Attended Community development forum
	p11	Voted in the 2016 local election
	p12	Not been asked for a bribe
Global life Satisfaction dimension	g1	Satisfaction: Life as a whole
	g2	Disagree tht no-one cares about me
	g3	Disagree that I cannot influence developments
	g4	The country is going in the right direction
Family Dimension	f1	Satisfied with marriage/relationship
	f2	Satisfaction: Family
	f3	Satisfaction: Time to do the things you want to do?
	f4	Satisfaction: Leisure time
	f5	Children skipped meal past year?
Community dimension	c1	People in your community can be trusted
	c2	Satisfaction: Friends
	c3	It is important to look after the environment
	c4	Have participated in an organisation/club event
Health Dimension	h1	Health status in the past 4 weeks
	h2	Health status prevent from doing daily work?
	h3	Health prevent from usual social activities?
	h4	Did not fail to get/look for healthcare
	h5	Personally covered by medical aid
Dwelling dimension	d1	Satisfaction: Dwelling
	d2	Satisfied with area where live
	d3	Dwelling structure made of bricks/concrete
	d4	Dwelling is owned
	d5	Not overcrowded
Connectivity Dimension	e1	Press is free to write/say what it likes
	e2	Highest level of education completed
	e3	Have telephone or cellphone
	e4	Household assets: Television
	e5	Household assets: Internet connection
Work Dimension	w1	Satisfaction: money available to respondent?
	w2	View as middle/upper class
	w3	Satisfaction: Standard of living
	w4	Satisfaction: Working conditions in your job
	w5	Employment Status
	w6	Total monthly income all household members
	w7	Not in debt
Security Dimension	s1	How safe walking in your area during the day?
	s2	How safe walking in your area after dark?
	s3	How safe do you feel at home?
	s4	Crime improved or got worse?
	s5	Have you been a victim of crime in the past year?
Additional Vars	r5	Government improved quality of life
	r4	8.22 - Communication with municipality

Right model (7-Factor QoL Index)

QoL index code	Reduced 7-factor solution	Variable	7-factor QoL Index
d3	0.800	Dwelling structure made of bricks or concrete	Infrastructure and services
i3	0.951	1.7 - Flush toilet facility	
i4	0.675	1.4 - Piped water source	
i5	0.827	1.9_Electricity supply	
i6	0.793	1.8 - Rubbish disposal	
e4	0.657	5.7_3 - Household assets: Television	
p6	0.848	6.6 - Satisfaction: National government	Government Satisfaction
p7	0.902	6.7 - Satisfaction: Provincial government	
p8	0.676	6.8 - Satisfaction: Local municipality	
r5	0.553	8.14 - Government improved quality of life	
p9	0.429	6.23 - Most government officials are doing their best?	
c4	0.443	10.1 - Participated in activities of any clubs	Political engagement
p10	0.711	6.4_3 - Attended Community development forum	
r4	0.548	8.22 - Communication with municipality	
p11	0.459	6.2 – Voted in the 2016 local election	
f2	0.658	7.4 - Satisfaction: Family	Life Satisfaction
f3	0.615	7.2 - Satisfaction: Time to do things you want to do?	
f4	0.664	7.7 - Satisfaction: Leisure time	
c2	0.547	7.5 - Satisfaction: Friends	
w3	0.360	7.6 - Satisfaction: Standard of living	
h2	0.899	11.16 - Health status prevent from doing daily work?	Health
h3	0.898	11.17 - Health prevent from usual social activities?	
h1	0.464	11.15 - Health status in the past 4 weeks	
h5	0.895	11.6 - Personally covered by medical aid	Socio-Economic Status
e2	0.662	12.1 - Highest level of education completed	
e5	0.547	5.7_5 - Household assets: Internet connection	
w5	0.491	Employment Status	
w6	0.850	12.10 - Total monthly income all household members	
s1	0.764	9.3 - How safe walking in your area during the day?	Safety
s2	0.556	9.4 - How safe walking in your area after dark?	
s3	0.620	9.5 - How safe do you feel at home?	
s4	0.534	9.1 - Crime improved or got worse?	
c1	0.461	6.43 - People in your community can be trusted	

Fig. 2 Dimensional comparison of the original (left) and factor-analysed (right) models

These questions are a mix of 'objective' items (e.g., type of housing) and 'subjective' items (e.g., reported satisfaction with various aspects of life). It is not documented whether a range of experts (Greco, 2019) or stakeholders (Hagerty et al., 2012) were consulted for this original specification, to avoid the subjectivity of the developer's perceptions (Booysen, 2002). The items are shown in the left-hand column of Fig. 2 (in Sect. 5.2), arranged into the original ten dimensions, and also with their levels of measurement in "Appendix".

In calculating this original index, the selected questions—comprising dichotomies, three-option questions, and five-point Likert scales—were firstly reduced to dichotomies. Missing or inapplicable values were set to 0, resulting in their being treated as negative

responses. The dichotomies, varying in number within in each dimension, were summed by dimension to yield the 10 dimension scores. These were scaled to run from 0 to 10, and then summed to yield the aggregate QoL index, in turn scaled to run from 0 to 10. Since the number of items per dimension ranged from 4 to 12, this approach was not unweighted, but amounted to an implicit, unequal weighting per item (Mikulić et al. 2015). An alternative approach was subsequently developed. This by-passed the dimensions and directly summed over all 54 dichotomies, and then scaled the total from 0 to 100. Use of this version has remained limited, and it will not be considered further. The questions, their dichotomisation, the arrangement into 10 dimensions, and the arithmetical compilation into dimension scores and an aggregated index were retained with negligible variation over the ten years from QoL I through to QoL V.

In preparation for the third survey, GCRO commissioned an exercise (Greyling, 2013) to consider alternative possible weightings of variables or dimensions during index compilation. This drew on quality of life theory and existing international indices to define 7 dimensions in advance, and select 27 of the 54 original indicator variables available from the QoL I survey. Using the QoL I data, it applied successive iterations of principal components analysis (PCA) to eliminate variables with the lowest communalities, and then applied a further PCA with varimax rotation to the remaining 15 variables. This approach did not—to anticipate the vocabulary of the next section—distinguish between reflective and formative indicators. The PCA indicated five dimensions, following the customary critera of eigenvalues greater than one, a scree plot, and interpretability (Tabachnik and Fidell 2007). Finally, the exercise followed the method of Nicoletti et al. (2000) in calculating each dimension's score by scaling to unity the sum of the squared loadings of its variables; and then in aggregating the final quality of life score as the sum of the dimension scores, after weighting them by their share of the total variance (i.e. their eigenvalues).

The aim of this exercise was to see whether it 'generated results that were significantly different' from GCRO's arithmetically compiled, 10-dimension index. According to Everatt (2017) 'the simple answer was "no".' The original approach was accordingly retained through to QoL V (De Kadt and Crowe-Petterson 2020).

2.3 Revisiting the QoL Index Variables, Dimensions and Calculations

In 2019 GCRO commissioned a review of the overall QoL survey programme, and a synthesis report (Orkin, 2020). An element of the review was that the four authors of this paper consider the data-handling and the steps of index construction used to date, and quantify how the results of a suitable statistically-based computation of the index might or might not differ from those of the prevailing approach. An alternative approach was developed, initially drawing on the much larger dataset of 24,889 responses available from QoL V of 2017/18 (GCRO 2018). As described in Sect. 4, we worked with the slightly increased selection of 58 quality of life variables, incorporated just 3 additional variables, corrected the processing of the missing or inapplicable values, and retained the ordinal Likert scaling of variables where it was used. Our application of exploratory polychoric factor analysis, and differently weighted compilations of dimensions and of the final index, were informed—but not dictated—by the indexing literature in the review (Orkin, 2020), including the approach previously recommended by the Greyling (2013) exercise and its subsequent iterations (Greyling and Tregenna 2017), which were still based on the QoL I data.

An additional objective of this work was to achieve a sound and well-precedented statistical approach that would be straightforwardly replicable in subsequent GCRO QoL analyses, in other South African provinces, and by lower and middle income countries in Africa and elsewhere. We believe that this further expectation is met by the stages of our analysis, yielding a more parsimonious set of 33 variables, re-arranged in 7 rather than 10 quality of life dimensions that are conceptually coherent and theoretically familiar. The stages are described in Sect. 4, Methodology.

Commensurately with the indexing process, Sect. 4 also details how we used CFA to confirm the robustness of the dimensional structure of this model against the two immediately preceding QoL IV and QoL III survey datasets (GCRO 2014; GCRO 2016), which provided similarly ample sample sizes. By contrast, we found that the original ten-dimensional, 58-variable model (Everatt, 2017) could not be fitted to these three datasets with a CFA at all.

As the literature makes clear, conceiving and constructing a composite quality of life index that is conceptually and statistically defensible, longitudinally comparable, and user-friendly, requires various informed methodological and theoretical choices; and, we would suggest, ultimately even philosophical standpoints. These arise from the complex nature of the observed societal 'reality' and the particular focus of quality of life, as well as the nature of quantitative indicators and their compilation first into dimensions and then a quality of life index (Maggino, 2014; Maggino & Zumbo, 2012).

Accordingly, in Sect. 3, Theory, we seek to define quality of life and attend briefly to the issue of whether it is appropriate to aggregate dimensions into a composite index. But, in particular, our approach must be defensible in the context of recent, vigorous debate in the quality of life literature about the nature and appropriate statistical treatment of 'reflective' against 'formative' indicators. Reflective indicators are exemplified in the scaling of items that are the observable manifestations of an unobservable latent construct such as depression, i.e. they can be said to be 'caused by' it. By contrast, formative indicators are observable items that jointly comprise a construct broader concept such as socio-economic status, i.e. 'cause' it; and by extension the construct may not be regarded as an 'underlying' construct at all. However, this too may be contested, which is where a philosophical debate will briefly become relevant. But we shall see that quality of life indices in practice have often treated formative as well as reflective indicators reflectively, and that to do so has even been argued to be empirically preferable. Without attending to the issue, Stiglitz et al. (2009) clearly expected that both formative and *prima facie* reflective items, and by extension dimensions, would be present in a quality of life survey; and this is often the case, especially in respondent- rather than country-level surveys.

In Sect. 3 we therefore defend a coherent route through selected contributions to this demanding debate at the different levels, while noting the wise view of Greco et al. (2019) that 'each choice made for the construction of a composite index appears to be "between the devil and the deep blue sea".' Therefore, we have accorded with his recommendation that 'robustness analysis should follow the construction of an index'.

Finally, Sect. 5, Results, compares results from our reworked approach with the original approach, to provide some revealing indicative findings. Section 6, Discussion, reflects on illustrative differentiations of the results, and limitations and implications of the investigation.

3 Theory: Definitions and Debates

3.1 The Concept of Quality of Life

As Stiglitz et al. (2009) allow, the term 'quality of life' is frequently used interchange-
ably with 'well-being'. According to Gasper (2010), 'well-being' is often used in psy-
chology or health literature to refer to experiences or circumstances as experienced by
an individual person, whereas 'quality of life' is more often used in sociology and social
policy to include the societal environment in which individuals find themselves. None-
theless, there is substantial overlap between the concepts, which extends to the meas-
ures and approaches used to quantify them. So, instruments for each tend to incorporate
variables from the other (Lourenço et al., 2019). We shall generally refer to quality of
life unless a source indicates otherwise.

On Glatzer's (2012) distinction, indicators that capture individuals' subjective expe-
riences or circumstances (e.g. 'satisfaction with your life as a whole' or 'feeling secure
in your neighbourhood') are said to measure quality of life subjectively. On the other
hand, quality of life is said to be measured objectively when individuals' living condi-
tions (e.g. housing, or environmental conditions) are assessed without asking them for
their opinion. (Respondents' personal reports of objective features, e.g. of their income,
are generally also classified as objective.)

The OECD (2020), for its country-level comparisons primarily using objective
administrative and official data, has recently defined 'well-being' by 14 dimensions.
Somewhat confusingly, it labels 11 of these dimensions as 'quality of life': work-life
balance, health, education, social connections, civic engagement and governance, envi-
ronmental quality, personal safety and subjective well-being. The other 3 dimensions
are 'material conditions': income and wealth, jobs and earnings, and housing. We may
take all 14 dimensions as a comprehensive ostensive definition of quality of life.

Several well-known quality of life surveys using objective or subjective indicators,
aggregated into differing numbers of dimensions, are mentioned in Sect. 3.2 below. The
individual-level GCRO QoL surveys contain both objective and subjective indicators as
reported by respondents—this is noted in Sect. 4.1.

3.2 Composite Indices of Quality of Life

Nardo et al. (2005) define a composite index as a single index derived from the process
of compiling individual variables based on a model designed to measure a multidimen-
sional concept. Similarly, Michalos et al. (2011) define a composite index of well-being
as a unidimensional index encapsulating a multidimensional conception of human well-
being. Such a composite index is valuable because it permits simple communication
about quality of life and its changes to an everyday audience, even though quality of life
itself should not be regarded as a unidimensional socio-economic phenomenon that can
be measured by a single variable (Mazziotta & Pareto, 2017).

Typically, a composite index is constructed in two broad steps: first by combin-
ing specifically operationalised variables—which can be put as questionnaire items
to respondents, or extracted from administrative data—into several intermediate-level
composite dimensions or constructs that reflect distinct aspects of quality of life (OECD,

2008); and then by aggregating the several dimensions into the single composite index (Mazziotta & Pareto, 2017).

The choice of whether or not to aggregate dimensions into a single index is debated in the literature (Greco et al., 2019; Saltelli, 2007). Proponents of stopping at the first step and retaining separate dimensions of quality of life argue that there is unacceptable 'information loss' (Lourenço et al., 2019) in the use of a single index, which may mask important changes in the dimensions. Instead, they advocate a 'dashboard' approach, by which the different dimensions of quality of life are individually monitored (Mazziotta & Pareto, 2017; OECD, 2008). Others (Saltelli, 2007) argue that this approach adds complexity and difficulty in communicating the results, especially to non-technical policy makers. Stakeholders, particularly those within government or the media, may expect and use a single composite index score, even as they may also be keen to engage with its component dimensions.

For example, in Italy, a newspaper published its composite quality of life index broken down by province, based on six dimensions of six variables each, and its trends have been widely scrutinised (Lun et al., 2006). Similarly, GCRO's previous composite QoL index results have been launched biennially in a high-profile event attended by the Premier of the Gauteng province. The results receive extensive media attention, for example in comparisons of the overall quality of life standing of the province's nine constituent municipalities (Hosken, 2018). Changes in municipal standing have been considered on the basis of each of the ten different constituent dimensions, as well as the overall quality of life score.

Accordingly, in re-working the previous GCRO QoL index this paper retains the overall two-step approach. Indicator variables are compiled statistically into different dimensions of quality of life, which can be presented and monitored on a 'dashboard'. But in addition, the constructs are statistically aggregated into a single composite index of quality of life.

3.3 Theoretical Framework, with Applicable Schema from Prior Cognate Studies

The OECD Handbook (2008) provides a checklist for the creation of a composite index: theoretical framework, data selection, imputation of missing data, normalization of variables, multivariate analysis, weighting and aggregation, sensitivity analysis, and identification of main drivers (The last two items are beyond the scope of this paper). Mazziotta and Pareto (2017) have digested similar broad guidelines from a review of several instances. The theoretical framework 'provides the basis for the selection and combination of variables into a meaningful composite indicator under a fitness-for-purpose principle', with the aim of getting 'a clear understanding and definition of the multidimensional phenomenon to be measured'. The remainder of this section, along with the next, explore approaches to establishing a theoretical framework to guide the creation of a composite index, and highlight how our approach differs.

Nine analogous studies were reassuring to different extents in their conception and number of indicators and dimensions, and their approach to aggregating dimension scores into a one-number index. Estes (2019) constructed the Weighted Index of Social Progress (WISP) based on forty indicators grouped into ten dimensions, namely education, health status, economic, demographic, environmental, social chaos, cultural diversity, and welfare effort. The statistical weights applied to the indicators and dimensions were derived from a two-stage principal component analysis and oblique factor analysis. The Australian Personal Well-being Index (Cummins et al., 2012) measures subjective well-being by averaging the level of satisfaction across seven dimensions of quality of life, namely standard

of living, health, achieving in life, personal relationships, personal safety, community-connectedness and future security. Stanojević and Benčina (2019) constructed a composite index of well-being based on three dimensions, i.e. health and education, political system and economy, by computing a weighted sum of indicators from factor analysis.

Ihsan and Aziz (2019) ranked districts in three provinces of Pakistan on the basis of quality of life, with a set of 31 variables grouped into seven dimensions. They used the percentage variance explained by each of the factors, the eigenvalues, as weights in the aggregation of the index. Verdugo et al. (2005) built on a polychoric correlation matrix because their 41 items were ordinal in scale. They validated their five-dimension structure with a confirmatory factor analysis enriched by fit statistics. Helmes et al. (1998) noticed that oblique rotation improved the fit of their factor structure to the data.

But the devil is in the detail. In the context of South Africa, in the latest of their three successive papers Greyling and Tregenna (2020) retained the classical approach from their earlier two papers (Greyling, 2013; Greyling and Tregenna 2017): of defining their sets of dimensions a priori, by comparison with previous international quality of life studies. The dimensions in these studies are summarised in Table 1 in Greyling and Tregenna (2020). Their chosen dimensions are housing and infrastructure, social relationships, socio-economic status (SES), health, safety, and lastly 'governance'. Then, again guided by the literature, they selected and assigned 33 variables to the six dimensions, out of the items used for quality of life in QoL I analysis; used EFA on each such dimension successively to check its unidimensionality; and finally validated the six-dimension model with CFA.

The point to carry forward is that their last dimension, 'governance' is incompletely defined and populated. A fuller understanding of this dimension is elaborated at length in Stiglitz et al. (2009) as 'governance *and voice*' [our italics]: the top-down emphasis of governance is importantly complemented (or countered!) by the bottom-up emphasis of citizen voice and active engagement. As stated (p. 177),

Political voice is an integral dimension of quality of life, having both intrinsic and instrumental worth. Intrinsically, the ability to participate as full citizens, to have a say in the framing of public policy, to be able to dissent without fear and to speak up against perceived wrongs—not only for oneself but also for others—are essential freedoms and capabilities. Political voice can be expressed both individually (such as by voting) and collectively (such as joining a protest rally).

Given the restricted emphasis of Greyling and Tregenna (2020) on governance, in choosing five variables for this dimension they include only one variable that, abstractly, addresses bottom-up voice and engagement: 'Politics is not a waste of time'. This is an item on political alienation drawn from a survey by Everatt and Orkin (1993). Their other four items are respondents' perceptions of top-down processes: 'Country is going in right direction', 'Elections were free and fair', 'Satisfaction with local government', and 'Judiciary is free'.

This seemingly minor conceptual attenuation has methodological and empirical implications. We find *empirically* that our approach, within a robust model, results in a seventh dimension, distinct from Greyling and Tregenna's (2020) largely top-down governance dimension of satisfaction with different levels of government. It comprises four bottom-up indicators of active political voice and engagement, close to that envisaged by Stiglitz et al. in the quotation above. As will be seen in Sect. 5, Results, the four indicators are: 'Participated in the activities of any clubs'; 'Attended community development forum'; 'Communicated with municipality'; and 'Voted in the 2016 local election'.

Methodologically, this points to the worth of our taking full advantage of the exploratory capacity of EFA, to both identify and populate dimensions, while drawing on the full set of

Table 1 Constituent variables in the seven factor models before and after indicator reduction

Labels	Indicators	Final 7 factor	7 factor	8 factor	9 factor	10 factor
Infrastructure and services						
d3	Brick/concrete dwelling	0.800	0.822	0.815	0.806	0.813
i3	Flush toilet facility	0.951	0.934	0.938	0.940	0.920
i4	Piped water source	0.675	0.655	0.657	0.662	0.660
i5	Electricity supply	0.827	0.836	0.823	0.828	0.820
i6	Rubbish removal	0.793	0.780	0.782	0.781	0.789
e4	Household assets: Television	0.657	0.648	0.642	0.640	0.645
d1	Satisfaction: dwelling		0.377	0.374	0.365	0.373
Socio-economic status						
h2	Personally covered by medical aid	0.895	0.842	0.838	0.829	0.836
e2	Highest level of education completed	0.662	0.670	0.665	0.656	0.662
e5	Household assets: internet connection	0.547	0.583	0.508	0.553	0.570
w5	Employment status	0.491	0.493	0.506	0.522	0.507
w6	Total monthly household income	0.850	0.802	0.799	0.799	0.782
f5	Children skipped no meals in past year		0.486	0.494	0.498	0.486
e3	Household assets: Cellphone or landline		0.390	0.381	0.376	0.379
w4	Satisfaction: working conditions in job		0.422	0.420	0.424	0.425
p12	Not asked to pay a bribe		− 0.361			− 0.351
Government satisfaction						
p6	Satisfaction: national government	0.848	0.815	0.811	0.824	0.825
p7	Satisfaction: provincial government	0.902	0.850	0.869	0.880	0.874
p8	Satisfaction: local municipality	0.676	0.664	0.671	0.675	0.669
r5	Government has improved quality of life	0.533	0.581	0.582	0.582	0.578
p9	Agree most government officials doing their best	0.429	0.451	0.449	0.425	0.433
Life satisfaction						
f2	Satisfaction: family	0.658	0.678	0.675	0.669	0.687
f3	Satisfaction: time to do things you want to do	0.615	0.603	0.602	0.592	0.605
f4	Satisfaction: leisure time	0.664	0.645	0.644	0.634	0.638
c2	Satisfaction: friends	0.547	0.530	0.535	0.527	0.527
w3	Satisfaction: standard of living	0.360	0.398	0.394	0.415	0.415
Safety						
s1	Feel safe walking in area during the day	0.764	0.735	0.783	0.707	0.763
s2	Feel safe walking in area after night	0.556	0.555	0.524	0.552	0.547
s3	Feel safe at home	0.620	0.623	0.650	0.613	0.636
s4	Crime situation improving	0.534	0.559	0.533	0.558	0.552
c1	Agree people in community can be trusted	0.461	0.467	0.449	0.469	0.464
s5	Not a victim of crime in past year		0.368		0.366	
Health status						
h2	Health does not prevent daily work	0.899	0.910	0.944	0.778	0.787
h3	Health does not prevent usual social activities	0.898	0.908	0.928	0.780	0.785
h1	Good health status in past 4 weeks	0.464	0.440	0.409	0.472	0.474

Table 1 (continued)

Labels	Indicators	Final 7 factor	7 factor	8 factor	9 factor	10 factor
Political engagement						
c4	Participated in activities of any club	0.443	0.412	0.367	0.384	0.400
p10	Attended community development forum	0.771	0.663	0.678	0.633	0.656
r4	Wants to hear from municipality	0.548	0.521	0.544	0.541	0.550
d4	Dwelling owned			0.464		
p11	Planning to vote in next election	0.459	0.499	0.569	0.543	0.522
Other						
p4	Disagree blacks/whites never really trust each other				0.919	
r1	Long travel time from home to frequent destination				0.791	0.918
e1	Agree press free to say/write what it likes					− 0.693
p5	Positive attitude to foreigners in Gauteng					− 0.737

The entries are standardized factor loadings. Additional columns illustratthe eight to ten factor models

available indicators (by the time of QoL V). By allowing the EFA to *generate* configurations of indicators and constituent variables, for empirically informed consideration, we let the relationships in the data—the painstakingly elicited responses of some 24,800 diverse Gauteng residents to 60 questions—revail on which indicators were to be selected, and how they congregate into dimensions. We believe that this is especially appropriate in very unevenly developed lower and middle income country contexts such as South Africa, where the patterns across multiple responses, and indeed the importance attached to the questions, may differ considerably in dense informal shack settlements or deep rural villages from those in formal suburban housing, for example. In this vein, within an overview of surveys in such contexts (Land et al., 2012), Shin (2008) stresses that despite Korea's transition 'from a low-income country into an economic powerhouse…Koreans neither interpret nor value democracy in the same way as Westerners do'. We believe that the manner and extent to which this variously manifests in the quality of life should be allowed, by the methodology, to be an empirical matter as far as possible.

Finally, there is the *philosophical debate* underlying the 'exploratory' approach to *generating* theory (or, more modestly, multidimensional quality of life models) out of data analysis, as against the 'hypothetico-deductivist' approach of testing prior theory (models) against the data, that we have now encountered. Indeed, operationally, they eventually converge, in that the exploratory approach will seek—in these instances—to validate its model by the same technique of CFA that is the end-point of the hypothetico-deductivism. This philosophical debate is beyond the scope of this paper, but we touch upon it in Sect. 3.5.

3.4 On the Nature of Quality of Life Indicators and Measurement Models: The Reflective/Formative Debate

The next, second item of the *OECD Handbook's* checklist, data selection, conceals difficult and far reaching presuppositions: are the variables under consideration, and the dimensions into which they will be somehow combined as measurement models, 'reflective' or 'formative'? As preliminarily characterised in Sects. 1 and 2, observable indicators are regarded

as reflective if they are functions of a latent construct or variable, 'whereby changes in the latent variable are reflected (i.e. manifested) in changes in the observable indicators' (Diamantopoulos & Siguaw, 2006). An example of reflective variables is the list of questionnaire items put to respondents to measure the underlying construct of depression.

By contrast, sets of variables are said to be formative when 'combined to form weighted linear composites intended to represent theoretically meaningful concepts' (Edwards, 2011). For example, changes in any of the observable variables comprising socio-economic status—education, income, etc.—will result in changes in the composite variable to which they contribute. Precisely speaking, the reflective variables associated with a latent construct comprise a *scale*, while formative indicators jointly comprise an *index*. The contrast between them is sometimes expressed causally (Diamantopoulos & Winklhofer, 2001): the 'direction of causality' runs from a latent construct to its reflective indicators, but in the other direction from formative indicators to the construct, typically a composite index. The former are correctly handled by factor analysis and traditional structural equation modelling, and the latter by principal components analysis (Tabachnik and Fidell 2007) or, more recently, by partial least squares structural equation modelling (Hair et al., 2021).

There is discussion as to whether this distinction should be made primarily on empirical or theoretical grounds. Both are favoured by Finn and Kayande (2005), Petter et al. (2007), and Coltman et al. (2008). Additional relevant considerations are that reflective indicators in a scale are expected to be highly correlated, and some may thus be abandoned without changing the construct; whereas formative indicators in an index may well not be highly correlated, so that abandoning one may substantially change the conception of the composite index. Indicators and their respective constructs constitute measurement models, whereas structural models establish relationships among constructs (Jarvis et al., 2003; Maggino, 2014). It is possible that that the measurement models for the extraction of dimensions may be conceived as reflective, and the structural model for the aggregation of dimension scores into an index may be conceived as formative (Jarvis, 2003), or vice versa; or they may match.

In treating formative indicators as reflective or vice versa, the measurement model is said to be mis-specified. But, as Diamantopoulos and Siguaw (2006) note, 'determining the nature and directionality of the relationship between a construct and a set of indicators can often be far from simple'. They show that the implications can be far-reaching. In an exercise comparing reflective and formative constructions of a dimension, they find that the number of indicators varies, and indeed that the indicators finally included are disparate.

In applying these careful distinctions to the original, stipulative GCRO approach, it appears initially (see Fig. 2) that the objective indicators constituting three of the dimensions were formative: in the socio-economic dimension, for example, the presence of piped water, flush toilet and electricity; in the dwelling dimension, whether the structure is made of bricks, and is owned; and in the connectivity dimension, highest education, and ownership of cellphone or television. The other seven dimensions appear to be more clearly reflective, containing items regarding satisfaction—with family, health, security, work etc., or overall.

On closer inspection, however, nearly all the ten stipulated dimensions are actually a mixture of reflective and formative items, to a greater or lesser extent: for example, there are seemingly reflective items of satisfaction within the infrastructure and dwelling dimensions, and conversely, formative elements of employment status and income in the work-satisfaction dimension. How should such combinations have best been handled? In effect, since the assignment of variables into dimensions as well as their combination into an index were originally stipulated by the developer and then treated

additively, both the measurement models and the 10-dimensional structural model were effectively formative.

The challenge, however, for our QoL index revision was how to proceed in regard to this reflective/formative issue. It will be seen in Sect. 4 below that, not unusually, we treat the relation of measures to dimensions, and also of dimensions to index, reflectively. As with the adoption of an appropriate theoretical perspective above, so with our selection and construal of variables: the route we adopted was necessarily pragmatic, but again has methodological support, precedent, and a philosophical foundation. *Pragmatically*, the given selection of 60 subjective and objective indicators we were revisiting was fortunately extensive; relevant in coverage, having been informed by the prior review (Jennings, 2012); and covering, as noted just above, a spread of putatively formative or reflective items.

Methodologically the overall reflective approach that we take on the reflective/formative issue accords with the position of Edwards (2011) and Simonetto (2012). They distinguish five objections to the use of formative indicators, which quite largely follow from the respective definitions elaborated above. The first is *dimensionality*: a reflective latent construct is unidimensional, in that its indicators are highly correlated, and therefore replaceable. But formative indicators represent distinct aspects of the index they comprise, with low correlation. This renders the index conceptually ambiguous. Secondly, *internal consistency* is a prerequisite for the indicators of a reflective construct, in that they correlate; but a problem for formative indicators, in that multicollinearity becomes problematic the more consistent they may be. Thirdly, a reflective measurement model is *identified* provided it has three or more indicators; but a formative construct needs 'to be inserted into a larger model' (Simonetto, 2012) involving one or two reflective indicators, which may well affect its loadings and its meaning. Fourthly, reflective measurement models incorporate *errors* of the indicators, whereas formative models implausibly assume that these are error free. Finally, the *construct validity* of a reflective measure is based on the extent of the loadings of its measures on the latent construct, as evidenced for instance by Cronbach's alpha. But such measures for formative constructs are altered by the additional indicators required for identification.

These advantages of the reflective over the formative approach may also be *empirically* sustained. As a counter to the experiment mentioned earlier of Diamantopoulos and Siguaw (2006), a full path-model study (von Stumm et al., 2013) measuring life-course pathways to psychological distress found that the formative model underestimated path coefficients, and there was difficulty in finding suitable identifier variables; whereas the reflective model was more sensitive, and had better fit.

We may also invoke *precedent*. Among the studies mentioned at the end of Sect. 3.3, Ihsan and Aziz (2019) notably use eigenvalues from their second-step factor analysis, i.e. of reflective dimensions, to weight the aggregation of their final quality of life score. And in South Africa Greyling and Tregenna (2020), considered above, moved in their successive analyses of the GCRO QoL I data from their earlier use of principal components analysis, consonant with formative measurement models, to using exploratory oblique factor analysis, i.e. presuming reflective indicators, to check the unidimensionality of their hypothesised dimensions; plus subsequent confirmation of the validity of the reflective dimensions using confirmatory factor analysis (CFA).

Finally, *philosophically*, this decision has a powerful implication. It makes easy sense with a psychological latent construct like depression that changes in the construct 'cause' its observable indicators to change. But surely the direction of causality is the other way around with SES? Not necessarily! This is where the standpoint of *critical realism*

becomes relevant. We briefly take this up in the next section, including its relation to the exploratory approach broached in the previous section.

On all these grounds, we believe it is defensible to resort to a reflective extraction both of dimensions and of the ensuing one-number index. In Sect. 4, Methodology we describe in Sect. 4.2 our resulting methodology of multivariate analysis. The other steps of the OECD checklist are taken up in the remainder of the section.

3.5 Philosophical Foundations of Competing Approaches

We noted earlier the remark of Greco et al. (2019) that 'each choice made for the construction of a composite index appears to be "between the devil and the deep blue sea".' In treading this perilous path, we defended our choice of starting with a boldly exploratory use of EFA, in Sect. 3.3. This entailed treating our dimensions and their constituent indicators as reflective, as argued in Sect. 3.4. As promised, the present section now signposts the competing standpoints in philosophy of science underpinning what we called the classical approach and our alternative—and indeed their rapprochement, if only in a final methodological step.

In Sect. 3.3 we noted that the 'classical' approach to compiling a composite multidimensional indicator applies, albeit in compacted form, the hypothetico-deductive philosophy (Popper [1935], 1959) of scientific theory testing. Having established their hypotheses—based on the prior literature or intuition—scientists deduce predictions that are testable against observable evidence. In this case, as we saw for Greyling and Tregenna (2020), the a priori configuration of six dimensions and their constituent indicators simultaneously provided the overall hypothesis and its testable implications. After the indicator assignments were separately tested for unidimensionality with EFA, the fit or otherwise of the overall CFA against the data, using standard criteria, provided the test of the configuration.

Our approach, by contrast, was informed by the equally venerable (e.g. Tukey, 1969) philosophy of exploratory enquiry. It basically asks the question, 'What is going on here?' (Behrens, 1997). This may be because of lack of relevant literature to inform prior hypotheses in a new field; or, as in our case, because we wished to allow for the possible effects of a context with highly uneven levels of development on indicator salience and dimension formation, rather than assume that influential quality of life measures primarily in the developed North would apply. The exploration begins with the data, for which we had the very large QoL V survey sample, including a generous set of 60 objective and subjective quality of life-related indicators. The exploratory enquirer fits tentative models with appropriate statistical techniques, in our case EFA, to supply configurations of dimensions with constituent indicators; and then seeks to improve their fit and interpretability by an iterative process of model re-specifications. For EFA this is constrained by a number of established criteria for a defensible number of dimensions.

In other words, exploratory enquiry is a mainly process of theory-generating. But to avoid over-fitting and Type I error, the enquirer expects (resorting to the hypothetico-deductive approach) to test the finalised model: on a 'hold-out' sub-sample, or more recently 'bootstrapping', or best of all a matching independent data-set. In our case, in addition to applying CFA on the modelling data, we had available the earlier QoL IV and III datasets, against which to check the model for configural invariance.

In practice, the CFA furnishes 'modification indices', which allow the hypothetico-deductivist some exploratory adjustments, provided they are minor, to 'save the hypothesis' (Lakatos, 1980). For the exploratory enquirer, as we saw, incorporating these adjustments

in their last model is merely a further final stage in exploratory model-building, following their CFA test. This momentary rapprochement does suggest that in the stage of building a composite index in this way, the scientist expects to move through phases of both explanatory inquiry and hypothetico-deductivism: 'hypothesis driven and data-driven science are not in competition with each other but are complementary and best carried out iteratively' (Kell & Oliver, 2004).

Finally, there is the matter of critical realism, raised in Sect. 3.3, where we drew on Edwards (2011) for the position by which seemingly formative indicators are seen as actually reflective of latent constructs existing in in the world. This position is a corollary of realist philosophy of science (Gorski, 2013). On this view, SES—or, perhaps less surprisingly—social classes, economies, ideologies, agency, and so on are real entities that have ontological existence separately from their manifestations, measures, and our thought about them. This view crucially meets 'the prerequisite for causality … that the variables involved refer to distinct entities'; whereas with formative constructs 'one variable is part of another, [and] then their association is a type of part-whole correspondence, not a causal relationship' (Edwards, 2011).

But realism does immediately pose the considerable challenge of researching the mediations relating theories and their empirical refractions to the ontology (Archer et al., 1999). A proposal particularly suitable to our approach, and our hopes for its policy impact, sees an adaptation of 'grounded theory' (Glaser & Strauss, 1967), a particular form of iterative exploratory enquiry, 'providing a method for critical realism' (Oliver, 2011).

4 Methodology

In the previous Theory section, we covered the first two points of the OECD checklist, theoretical framework and data selection, within the context of revisiting GCRO's QoL index. We are positioned to take forward a two-step reflective analysis, drawing on 61 indicator variables listed in "Appendix". The original 58 categorical variables were augmented with three new, longitudinally available, user-relevant variables from QoLs III to V (denoted r1, r4 and r5 in "Appendix"); and one variable, the participant's rating of their overall quality of life, was removed to serve as a possible reference indicator. This left a total of 60 variables which were considered in our analysis.

As before, there are choices to be made among possible techniques at each juncture (Greco et al., 2019; Maggino & Zumbo, 2012). In this section we give a methodological account of how we handled the next five points of the OECD checklist: comparability of variables, imputation of missing data, multivariate analysis, weighting and aggregation, and our validation exercise in lieu of a sensitivity analysis. We shall touch on the last point, identification of main drivers, in Sect. 6, following the Results section.

4.1 Comparability of Variables

All the variables in the questionnaire were categorical: either dichotomies (0 or 1), or three- or five-point Likert scales. Variable coding was reversed where necessary so that higher values indicated increasingly positive responses. The 3 and 5-point scales were re-set to start from 0, like the dichotomies. In contrast with the original indexing, they were not dichotomised, in order to respect the nuances in the data. This work was conducted using SPSS v27.

4.2 Missing Values

Structurally missing values (e.g. missing values arising from 'inapplicable' questions such as a question on marriage satisfaction for single people) were set to the neutral mid-point for Likert scales. Values which could be classed as missing at random (i.e. where respondents said 'Don't know' or declined to answer, or the fieldworker simply overlooked a question) were imputed using the 'missForest' software package in R (Stekhoven & Bühlmann, 2012), a non-parametric multiple imputation method suitable for categorical data.

The process of imputing the variables with missing values was desirable to avoid pairwise or listwise deletion of variables during computation of the correlation matrix for the EFA. It also prevented changes in sample size as variables were removed during EFA. For example, the variable w6, 'Total monthly household income', had missing values for approximately 35% of observations, meaning that listwise deletion would have resulted in a substantial reduction in sample size.

4.3 Multivariate Analysis

4.3.1 Exploratory Factor Analysis (EFA)

As discussed in Sect. 3.3, our two-stage approach to the use of factor analysis was guided jointly by our preference – based on theoretical, empirical and philosophical grounds—for a broad exploratory approach to both the selection of indicators and the population of factors (dimensions); and for reflective rather than formative measurement as discussed in Sect. 3.4. The EFA technique groups correlated observed variables, or 'indicators', into distinct unobserved dimensions which are also referred to as latent constructs (Stanojević & Benčina, 2019). The obtained latent constructs are then assumed to 'cause' the observed variables in the respective dimensions; equivalently, the observed variables 'reflect' the respective latent constructs. The correlation obtained between a variable and its causal factor is referred to as factor loading (Cooper, 1983). Ideally, each indicator loads strongly on only one dimension. However, dimensions are likely to be correlated in practice, as they are in the social world to varying extents. In factor analysis, the particular cross-loadings can be greatly reduced by invoking *oblique* rotation of the dimensions (Plucker, 2003).This makes factor analysis the method of choice in the construction of a composite QoL index.

By contrast, principal components analysis (PCA) seeks, ideally, a single principal component that is the linear combination of variables that explains the most variance in the data. It is consonant with the conception of lightly correlated indicators formatively contributing to an outcome, typically a multivariable index. However, more than one principal component (dimension) may be extracted in PCA, in which case the dimensions are by definition *orthogonal* and uncorrelated (Howard, 2016). This orthogonality is retained by the choice of rotation, e.g.varimax, that may be undertaken to ease interpretation.

In EFA, the characterisation of the dimensions is suggested by the one or two highest loading variables, and the process of retaining variables, including the choice of a minimum cut-off, is then informed by how variables load and also how they contribute to the meaning of the dimension under which they fall. For these reasons our choice departed from the use of PCA by, for example, Ihsan and Aziz (2019), as well as Greyling (2013) and her precedent in Nicoletti et al. (2000). But our comprehensive exploratory application of EFA also departed from the later approach of Greyling and Tregenna (2020), who

confined their use of EFA to confirm or refine the unidimensionality of the dimensions and contents that they had selected previously, on the basis of prior specifications of dimensions in use in, primarily, the Global North.

To accommodate the categorical nature of variables, our EFA used a polychoric correlation matrix, like Verdugo et al. (2005). The most frequently used methods of extracting the factors in a factor analysis include the maximum likelihood method and principal axis factoring, which are suited for continuous (i.e. interval and ratio) variables. Given that the data employed in this study are categorical (i.e. binary and ordinal), weighted least squares (WLS) was the appropriate method of factor extraction. Oblique rotation was applied to allow for the correlation between dimensions, using the oblimin variant.

4.3.2 The Number of Dimensions

Four criteria were considered to determine the optimal number of dimensions, or factors (Afifi et al., 2019). Firstly, on the Kaiser criterion, dimensions with eigenvalues equal to or greater than one are retained. Secondly, parallel analysis compares eigenvalues obtained through factor analysis on the actual dataset and on a similar dataset populated with random numbers. Dimensions with higher eigenvalues from the actual dataset compared to their equivalent from the randomised dataset are then retained. The third criterion examines a scree plot for the 'knee'. The fourth criterion is the conceptual coherence and interpretability of dimensions, given the distribution of variables among them.

4.3.3 Rationalising Indicators Per Dimension

In the model with the chosen number of dimensions, all variables with absolute factor loadings greater than or equal to 0.35 were considered for retention (Howard, 2016). Then, following Kenny's (2016) recommendations for desirable numbers of indicators per dimension, all variables with loadings greater or equal to 0.5 were retained. Where there were fewer than five such variables in a dimension, additional variables with factor loadings less than 0.5 but greater than or equal to 0.35 were additionally retained, from highest factor loading to lowest, until the dimension included five variables, or until all variables with factor loadings of 0.35 or above were included.

The factor analysis yielded fit statistics and also an eigenvalue for each factor (dimension), to be used in weighting the dimension scores in the second stage, the aggregation of the index. The established criteria for an 'excellent' model (Brown, 2015) are root mean square error of approximation (RMSEA) < 0.05 and root mean square residual (RMSR) < 0.05; and comparative fit index (CFI) and Tucker-Lewis index > 0.95. However, earlier, less stringent criteria are taken as 'good' for more diverse, non-psychological models: < 0.06 and > 0.90, respectively. Our fit statistics were assessed against these criteria.

4.3.4 Testing the Dimensions' Construct Validity

After factor analysis has identified the dimensions and the contributing variables have been established, Cronbach's alpha is useful in testing the separate internal consistency of each dimension (Tavakol & Dennick, 2011), i.e. the extent to which the constituent variables measure the same concept. Cronbach's alpha runs from 0 to 1, and a value of 0.7 or more is typically considered acceptable for uses such as testing a relatively homogeneous psychological

scale for a latent variable. However, given the great diversity of indicators, we allowed for a marginal zone down to 0.6.

4.3.5 Checking EFA Model Fit with Confirmatory Factor Analysis (CFA)

Following the EFA establishment of dimensions and the refining of their indicator selections, the overall measurement model was tested by CFA, initially against the same GCRO QoL V data (Greyling & Tregenna, 2020), and then subsequently using the data from earlier iterations of the GCRO QoL survey (see Sect. 4.5). CFA is a type of structural equation modeling which requires the analyst to specify the model's parameters a priori based on prior knowledge about the sample data (Brown, 2015). These parameters are estimated to predict a variance–covariance matrix which approximates the variance–covariance matrix of the sample data (Brown, 2015). The aim is to confirm that there is construct validity. Different levels of invariance can be tested across samples and also over time (Liu et al., 2017). This study sought to check 'configural invariance', which tests the assumption that the pattern of indicators among the dimensions has remained the same over successive QoL surveys.

We implemented all of our EFA and CFA in R.

4.4 Weighting and Aggregation

The composite QoL index was computed from the overall model in the following three steps:

4.4.1 Scaling of the Variables

All non-dichotomous variables were re-scaled to range from 0 and 1. For example, the variable 'Total monthly household income' (w6 in "Appendix"), a Likert scale ranging from 0 to 4 (after re-setting its start value to 0), was divided by 4. As monotonic transformations of variables, re-scaling did not alter the distribution of the original dataset.

4.4.2 Computation of the Weighted Dimension Scores

Each dimension score was calculated by weighting each of its transformed individual variables with the factor loading of that variable, then summing, and finally re-scaling the dimension total to run from 0 to 1.

In mathematical notation, the computation of dimension score D_s is given by the expression (1) below.

$$D_s = \frac{1}{\sum_{i=1}^{n_s} f_{si}} \sum_{i=1}^{n_s} I_{si} f_{si} \tag{1}$$

where I_{si} = the value of indicator i scaled to 0–1, f_{si} = the factor loading of indicator i in dimension s, n_s = total number of variables within dimension D_s.

4.4.3 Aggregation of the Dimension Scores into the QoL Index

The overall QoL index was calculated by weighting each of the dimension scores with its eigenvalue, then summing the weighted dimensions scores, and finally re-scaling the total to run from 0 to 100.

In mathematical notation, the computation of the overall index score is provided by the expression (2) below:

$$\text{QoLindex} = \frac{100}{\sum_{s=1}^{m} E_s} \sum_{s=1}^{m} D_s E_s \tag{2}$$

E_s = eigenvalue of dimension s, m = total number of dimensions (in this case m = 7). The use of the value 100 as the numerator in this formula serves to scale the score to run from 0 to 100. The calculations outlined above were completed in SPSS.

4.5 Validation Exercise

The overall model validated by CFA on the QoL V data was then checked for 'longitudinal configural invariance' (Byrne, 2012) against the datasets for the two preceding GCRO QoL surveys—IV (2015/16) and III (2013/14). This checks whether the number of factors and the location of indicators is retained. The next level of invariance, 'construct-level metric invariance', of factor loadings and thus of the respective dimension scores, was not tested. Such changes, and the resulting change in the overall quality of life score, were expected; indeed, the examining thereof was much of the purpose of repeating the QoLs. Accordingly, in the runs against QoLs III and IV, parameters were freely estimated each time. Given the size of the model, minor adjustments were anticipated for convergence to be achieved. Consequently, we were able to calculate Index scores for QoLs III and IV using the model that was generated using QoL V.

5 Results

5.1 Modelling of QoL Indicators and Dimensions

5.1.1 Original EFA on 60 Variables

EFA with oblimin oblique rotation was applied to the re-coded data with imputed missing values for the 60-variable dataset for QoL V listed in "Appendix".

5.1.2 Number of Factors

The four criteria for likely numbers of factors were examined. The Kaiser criterion, retaining dimensions with eigenvalues equal to or greater than one, supported a seven-dimension model. Parallel analysis, retaining dimensions with higher eigenvalues than their equivalent from the randomised dataset, preferred 16 dimensions. The 'knee' in a scree plot suggested seven dimensions.

Regarding conceptual coherence, models for 8, 9 and 10 factors were examined in case this extended insight. But this mainly had the effect of adding low-loading variables to the robust 7 factor solution. Despite some scepticism in the literature about the Kaiser criterion (Howard, 2016), and given the corroborating result from the scree plot, 7 dimensions were tentatively adopted. They are shown in the column of Table 1 marked '7 factor', with the standardised loadings of their component indicators.

Table 2 Eigenvalues of the dimension in 7-dimensional EFA with rationalised indicators

Dimension	Eigen value
Infrastructure and services*	3.801
Socio-economic status*	2.695
Government satisfaction*	2.596
Health status	1.936
Safety*	1.811
Life satisfaction*	1.735
Political engagement	1.365

*Five dimensions, asterisked, have at least four indicators with loadings > 0.5. Health status has only 3 indicators, 2 of them > 0.8, while Political engagement has 2 of its 4 indicators > 0.5. No indicators < 0.35 were retained

Table 3 CFA fit statistics for the factor-analysed 33-variable and 7-dimensional model, for the three most recent iterations of the GCRO QoL survey

GCRO QoL survey	RMSEA	CFI	RMSR
QoL V (2017/18)	0.045	0.969	0.058
QoL IV (2015/16)	0.043	0.985	0.058
QoL III (2013/14)	0.044	0.984	0.059
Excellent fit criterion	< 0.05	> 0.950	< 0.05

Application of Kenny's (2016) recommendations, mentioned above, improved the parsimony by retaining all variables with loadings > 0.5, but augmenting factors with fewer than five such loadings with additional variables with loadings > 0.35. Following this process, a set of 33 variables remained.

5.1.3 Reduced EFA on 33 Variables

A follow-up EFA specifying seven dimensions was run on this reduced set of variables, to provide the factor loadings of the indicators on their respective dimensions that would be used in the stage 1 computations of the dimension scores; and also the eigenvalues for use as weights on those scores in the stage 2 computation of the overall QoL index score. The final dimensions and factor loadings are shown in Table 1, in the column 'Final 7' (Table 2). The eigenvalues for each dimension are provided in Table 2.

On the characterisation of Brown (2015), the fit statistics of this model were 'good' for RMSEA, 0.055, and excellent for RMSR, 0.02. Its TLI, 0.886, was a little short of the 'good' value of 0.9. This may be considered adequate in a large model with conceptually varied dimensions. (In Table 3 below, it will be seen that the CFA conducted with this model had excellent RMSEA and TLI, but only good RMSR, 0.058).

When the anomalous parallel analysis test was checked against the reduced set of 33 variables, it now suggested 8 factors. But it is seen in Table 1 that an 8th factor added only one variable 'Dwelling owned', rather surprisingly in the 'Political Engagement' dimension. Hence, the 7 factor solution remained optimal.

5.1.4 Cronbach's Alpha

Application of Cronbach's alpha to each of the 7 dimensions shows that, in this reduced 33-variable solution, three dimensions have alpha > 0.7, and three are > 0.6, our suggested criterion of internal consistency for such a variegated model. Only the 7th dimension, political engagement, is < 0.6, reflecting that two of its four factor loadings are < 0.5 but > 0.35.

5.1.5 Testing Model Fit with Confirmatory Factor Analysis (CFA)

Following development of the 33-variable, 7-factor model, this was then tested against the same GCRO QoL V data using CFA. Results for this model are presented in the top line of Table 3.

The table also shows the fit results of a test for overall configural invariance (Byrne, 2012), i.e. whether the number of factors and their respective indicators from QoL V fitted the datasets from QoL IV and QoL III. For QoL IV, only one variable among the 33— indicated by modification indices to be p10, 'Attended community development forum'— had to be dropped for convergence to be achieved. In QoL III, variable h1, 'Health status in the past 4 weeks' had not been included in the questionnaire. Given these small variations, Table 3 shows that according to RMSEA and CFI the fit of the model was 'excellent' on the two earlier data sets, and 'good' on RMSR. This licensed the absolute comparison across the scores of the Index as calculated using data from all three survey iterations.

5.1.6 Oblique Rotation

The need for oblique (oblimin) rotation in the EFAs is confirmed in the co-variations among dimensions obtained from the CFA, shown in Table 4. All the significances were $p < 0.001$ except for that between Political Engagement and Government Satisfaction, $p = 0.063$, and between Infrastructure and services and Health, $p = 0.165$.

5.2 Comparison of Configurations Between the Original and Factor-Analysed Approaches

The differing composition of the dimensions, between the original, 10-dimension, 60 variable model and the factor-analysed, 7-dimension, 33 variable model is shown in Fig. 2.

Obviously, the numbers of variables in cognate dimensions will mostly differ in two such different models. But the placing of variables is sometimes revealingly different, showing how the use of the EFAs exposes conceptually plausible patterning in the data that undercuts the original, stipulative approach. These shifts are taken up in the Discussion section.

5.3 Comparison of Index Scores Between the Original and Factor-Analysed Approaches

Aggregate QoL index scores for various demographic groups using the original 10-dimensional approach (Everatt, 2017) and the 7-dimension factor-analysed

Table 4 Covariance matrix among the 7 factors in the reduced factor model

	Infrastructure and services	Socio-economic status	Government satisfaction	Life satisfaction	Safety	Health	Political engagement
Infrastructure and services	0.777	0.423	0.186	0.134	0.119	0.006	0.034
Socio-economic status		0.771	0.146	0.149	0.167	0.057	−0.037
Government satisfaction			0.659	0.136	0.149	0.016	0.007
Life satisfaction				0.293	0.103	0.039	0.019
Safety					0.470	0.061	−0.028
Health						0.238	−0.026
Political engagement							0.187

All entries significant at $p < 0.001$ except Political Engagement and Government Satisfaction, $p = 0.063$ and Infrastructure and services and Health, $p = 0.165$

Table 5 Comparisons by gender and race of differences in overall QoL index between the original and the new, factor-analysed approaches, for QoLs III-V

Group	QoL III (2013/14)		QoL IV (2015/16)		QoL V (2017/18)	
	Original	New	Original	New	Original	New
Male	61.531	61.914	61.856	63.076	63.277	63.998
Female	60.518	60.835	62.084	63.527	62.654	63.869
African	58.953	59.502	59.781	61.119	60.840	61.734
Coloured	61.766	62.443	63.539	65.045	64.830	67.272
Indian/Asian	66.702	66.410	69.136	70.218	69.483	71.776
White	69.514	68.959	70.450	71.732	72.935	73.662
Overall	61.030	61.380	61.969	63.299	62.969	63.934

approaches are compared by population group and gender in Table 5. (The original QoL index scores were historically presented as scores out of ten. For ease of comparison, these have been scaled to scores out of 100.)

At the aggregated QoL index level across large numbers of variables (58 and 33 respectively), it is not surprising that the differences between the original and new indices are numerically slight. The trends—slightly differing improvements of overall quality of life—are sustained. However, when the scores are broken down by municipality, differences over time emerge which may be important in considering their relative progress. The differences are most clearly seen in terms of rankings of the municipalities, as in Table 6.

Most of the changes between the original and factor-analysed methods are by one rank position. In the two earlier QoLs there are some changes by two positions, e.g. in QoL IV, Merafong City from 7 to 5th. The rankings of Ekurhuleni are different for the two methods in all three survey iterations, moving from 3rd to 2nd to 4th over the last three QoLs, rather than 5th to 3rd to 3rd. Only the City of Tshwane's rankings were unaffected by the choice of method: a deterioration from 1st to 4th to 6th. These rankings are graphed, for the new approach, in the Discussion section.

## 6 Discussion of Results, Philosophical Foundations, and Limitations

This section consists of three parts. In the first part we consider some selected salient results of the revised QoL index based on our all-reflective, 2-stage, weighted, EFA using the QoL V data, and subsequently validated through CFA on earlier datasets. We make comparison between the original GCRO approach and our new analysis. This also allows us to point to selected 'drivers' for future research, our last duty to the schema from the OECD Handbook (2008).

In the second part, as promised in Sects. 3.3 and 3.4, we contrast our methodological sequence with that of a closely relevant and recent South African study using the QoL I dataset, by Greyling and Tregenna (2020). We indicate the consequent, marked difference in dimensions that arise empirically. Finally, in the third part we touch on some limitations and strengths of our analysis.

Table 6 Rankings of municipalities over time, for the original and factor-analysed indices

	QoL III		QoL IV		QoL V			
	Original rank	New rank	Original rank	New rank	Original rank	New rank	Original index	New index
City of Jo'burg	2	2	2	1	4	3	63.44	65.05
City of Tshwane	1	1	4	4	6	6	62.50	62.77
City of Ekurhuleni	5	3	3	2	3	4	63.44	64.27
Emfuleni	8	7	8	7	9	9	60.03	60.82
Lesedi	4	5	9	9	2	2	64.15	65.74
Merafong	6	6	7	5	8	7	60.76	61.93
Midvaal	7	8	6	6	1	1	66.64	66.04
Mogale City	3	4	1	3	5	5	62.84	63.87
Rand West	9	9	5	8	7	8	61.46	61.30
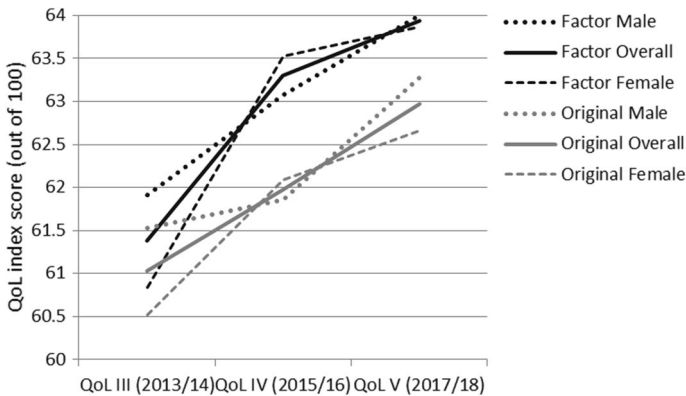
Fig. 3 Differences between the original and revised ('factor') QoL III–V scores, by sex

6.1 Salient Results

Beforehand, we may recall the seven dimensions of our model, drawn on 33 of the 60 selected quality of life indicators available in GCRO's QoL V (2017/18) dataset. The distribution of indicators in the original model and in our factor model was shown in Fig. 2. It was noted there that using factor analysis to determine the actual import of several indicators from respondents' replies through factor analysis revealed notable differences from the assignment of indicators to dimensions in the original model by the developer. 'Having medical aid' was a telling example, actually functioning in the EFA-based model as the strongest indicator in the SES dimension (because of its relation to formal better-paid employment), rather than in Health as allocated in the original model. This poses a challenge for the theory- and literature-based approach championed by the OECD (2008), to which we return below.

The dimensions (and rounded eigenvalues) of the revised model, from Table 2, are: Infrastructure and Services (3.8), Government Satisfaction (2.6), Health Status (1.9), Safety (1.8), Life Satisfaction (1.7) and Political Engagement (1.4). All but Health Status have four or more indicators, while all but Political Engagement have nearly all their indicators loadings > 0.5. That the highest eigenvalue is nearly three times the lowest is a reminder that applying them to weight the dimension scores in aggregating them (like Ihsan & Aziz, 2019) is a substantial adjustment to not weighting them, as in the original approach. Of the latter approach, Chowdhury and Squire (2006) remark that.

This is 'obviously convenient but also universally considered to be wrong.' Further, against the objection that choosing one over another weighting approach is subjective, Greco et al. (2019) note that not weighting is indeed also subjective.

In Table 3 in Sect. 5.1 we reported empirical evidence based on CFA validity checks across survey iterations, showing configural invariance of dimensions and their indicator content over time. Consequently, we are encouraged to examine trends in the revised QoL index scores over time. As the original index was extensively used for comparison over time, it is appropriate to explore the variation in the longitudinal trends derived from use of each model. For convenience, results first presented in Sect. 5 are illustrated here with graphs. The solid lines in Fig. 3, drawn from Table 5, show the difference over time between the original and the factor-analysed *overall* QoL index scores. Given the
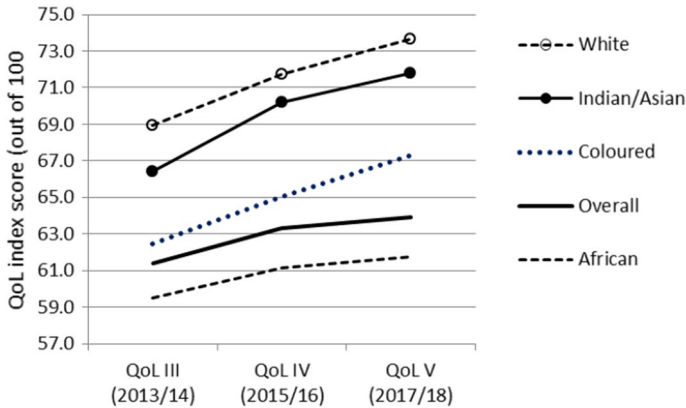
Fig. 4 Factor index scores for QoLs III–V, by race

aggregation over a large number of indicators common to both, it is to be expected that the differences at overall index level are small: at most 1.7 points at QoL IV. But these small differences are noteworthy given the very large samples. Figure 3 also shows the differences by sex over time.

On the original index, shown by the lower set of three lines, it appears that females' quality of life, from being 1 point behind males' at QoL III, had risen to slightly exceed that of males by QoL IV, and then fell back by QoL V. However, according to the new factor-based index, shown by the upper set of three lines, females' quality of life increased more sharply from QoL III to IV by 2.8 points, clearly to exceed that of males; so that, although it fell back, it still equalled that of males in QoL V. Further research could show which dimensions were responsible for the sharp initial improvement—a matter of considerable relevance in South Africa, where, for example, representation by sex has been legislated at all levels of government (Vyas-Doorgapersad & Bangani, 2020).

Thus sex is one 'driver' (OECD Handbook 2008) of index differences, albeit not a strong one. In post-apartheid South Africa, unsurprisingly, another driver is race. Using the factor index, Fig. 4 reflects the different levels, and different rates of improvement over time, among the four main population groups. White people have the highest average score, while Black African people have the lowest. Disturbingly, the overall quality of life gap between them is also still widening over time. (This was also evident in the original index.) Even though some QoL indicators show improved satisfaction with national government, others such as sanitation and health care have not, and some have turned downwards. The interplay between indicators suggesting progress, and those suggesting growing challenges, needs to be differentiated and further examined over time.

Another evident driver is the separate municipalities' reported organizational performance. Table 6 above, not graphed, showed how municipal rank orders differed at each point in time according to original and factor index scores. There were five such changes among the nine municipalities in QoL III, six in QoL IV, and two in QoL 5. Most were changes of one place in the ranking, but in a few cases there were changes of two places, and in one instance three. This positioning is accorded prominence by the Premier of Gauteng in response to QoL survey launches, and the extensive ensuing newspaper publicity (Sunday Times, 2018)—much as with analogous rankings in Italy
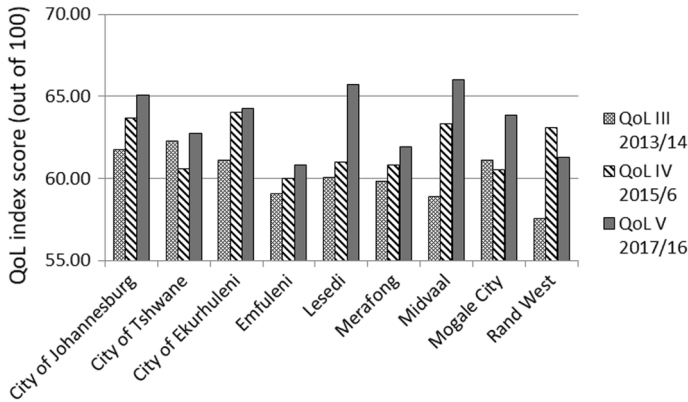
Fig. 5 Municipalities' QoL index scores at QoLs III–V

(Lun et al., 2006). The empirically selected and validated configuration in the factor index will help to obviate debate about changes in rank.

Changes over time in municipalities' seven scores for the constituent dimensions can also encourage investigations of the reasons for noteworthy changes, as seen in Fig. 5. The changes are quite striking. Notably, the initially highest index scores of Johannesburg and Tshwane at QoL III were exceeded just four years later, by the time of QoL V, by very rapid increases by Midvaal, from a low base. From a better base Lesedi, Emfuleni and Merafong also improved, but more gradually, and remained comparatively low. Rand West fell back at QoL V; and Tshwane actually fell back at QoL IV, and failed fully to recover by QoL V. This regress in Tshwane in 2016 evidently reflected citizen reactions to the widespread rioting and looting in Tshwane following the imposition by the ruling party of an unpopular executive mayor. (Maverick, 2016).

Municipal management, and monitoring by the provincial government, can be informed in even greater detail by the changes at indicator and dimensional level—for instance, respondents' assessment of the objective provision of services in the Infrastructure dimension, and their subjective satisfaction with the municipality in the dimension of Satisfaction with Government. These two dimensions were seen to have the second highest covariance, in Table 4.

6.2 Limitations

This study escapes some limitations common in the literature, such as those of the representivity and adequate size of the sample. Nonetheless, it has a number of limitations. The first of these is that the 60 candidate indicators listed in "Appendix", and the approach taken to their measurement, were a given—for reasons of economy and comparability. Sound practice would have reviewed their suitability. But the list was extensive, and informed by a review of previous developmental surveys. Certain objective variables, for example related to housing, were perhaps under-represented. Unfortunately, inclusion of new variables in future survey iterations, and incorporation into an index, would require adjustments to weighting and aggregation, in turn invalidating direct comparison with index scores from earlier survey iterations.

Secondly, because of the original mis-coding of variables, and the smaller earlier samples, recoding for use in the extra CFAs was taken back only to QoLs IV and III. However, before that the area sampling was less fastidious (Orkin, 2020) so this is perhaps as well. Thirdly, the successive samples are cross-sectional at the individual level rather than a panel, which would be prohibitively expensive on the necessary scale. But using the largely unchanged wards and municipalities (529 across the nine municipalities) as the unit of analysis, multi-level statistical analysis of trends would be entirely appropriate. In addition, more referencing with external variables to understand trends would be revealing: for example, with the Auditor General's (2020) five levels of annual audit outcomes.

Against these limitations may be set the evident utility of the index: the great attention to the results by the province's Premier, municipal governments, and residents; and as a result the many requests by municipal service-delivery departments, for example, for special analyses with which to respond to the Premier's queries and criticisms.

7 Concluding Remarks

Our proximate purpose, in undertaking a statistically based reworking of the GCRO's original QoL index, was to correct the misleading data-coding, discard extraneous indicators that are costly to collect, and replace the ad hoc indicator assignments and unweighted original aggregations. Using the QoL V data we were able to provide a parsimonious model—inally consisting of 33 indicators with appropriately weighted aggregations into 7 dimensions and a single QoL score—and establish its configural invariance across the two prior datasets. Additionally, this model may be updated in future, if a CFA fits well and survives validation, by obtaining from it revised weights for dimension scores and the single QoL score.

The indicative results are promisingly revealing, and can prompt a research programme within the GCRO and beyond: for instance, into drivers of the index and its dimensions, as well as dis-aggregations—subjective as well as objective—that will be informative for monitoring and planning.

Our distal purpose was to shape a simple but defensible methodological sequence, based entirely on straightforward recoding of variables, imputation of missing values, and some runs of EFA before use of CFA for validation. All techniques are available in any widely known statistical package, and could easily be replicated: in other provinces of South Africa, in fellow lower and middle income countries in Africa, and beyond.

This approach did prompt some penetrating enquiries by the journal's reviewers. Answering these required a brisk journey through methodological and finally philosophical questions. Our factor-analytic approach is overtly exploratory, which may not be comfortable to people schooled in prior hypothesizing and subsequent tests. And it in turn presumes that indicators (and scores for the aggregation of dimensions) are all reflective rather than formative, even for indicators such as SES. This posits a known but perhaps unusual philosophical understanding of what one is measuring.

A major benefit of the approach, alongside its coherent statistical simplicity, is its sensitivity to local context. This is highlighted by the discovery in our particular milieu of a separate seventh dimension of bottom-up political engagement or 'voice' that complements the dimension of satisfaction with top-down governance, and confirms the importance accorded to such measures by Stiglitz et al. (2009).

Appendix

The QoL variables and their measures from the original analysis, ordered in the original ten dimensions.

Variable	Label	Recoded/rescaled as likert scale or binary	Missing status
Has been an improvement in the community	i1	Likert scale (1 to 3/0 to 2)	No missing value
Water is usually/always clean	i2	Likert scale (1 to 5/0 to 4)	No missing value
Flush toilet facility	i3	Binary (0 to 1)	No missing value
Piped water source	i4	Binary (0 to 1)	No missing value
Electricity supply	i5	Binary (0 to 1)	No missing value
Rubbish disposal	i6	Binary (0 to 1)	No missing value
Have not had water/electricity cut off or been evicted	i7	Binary (0 to 1)	No missing value
Disagree that politics is a waste of time	p1	Likert scale (1 to 5/0 to 4)	No missing value
Agree that elections were/will be free and fair	p2	Likert scale (1 to 5/ 0 to 4)	No missing value
Agree that judiciary free from government influence	p3	Likert scale (1 to 5/0 to 4)	No missing value
Disagree blacks and whites will never trust each other	p4	Likert scale (1 to 5/0 to 4)	No missing value
Believe foreigners should be allowed to stay	p5	Likert scale (1 to 3/0 to 2)	No missing value
Satisfaction: National government	p6	Likert scale (1 to 5/0 to 4)	No missing value
Satisfaction: Provincial government	p7	Likert scale (1 to 5/0 to 4)	No missing value
Satisfaction: Local municipality	p8	Likert scale (1 to 5/0 to 4)	No missing value
Agree most government officials doing their best	p9	Binary (0 to 1)	Option 2 (Never interact with government officials) set as missing at random
Attended community development forum	p10	Binary (0 to 1)	No missing value
Planning to vote in next election	p11	Binary (0 to 1)	No missing value
Not been asked for a bribe	p12	Binary (0 to 1)	No missing value
Satisfaction: Life as a whole	g1*	Likert scale (1 to 5/0 to 4)	No missing value
Disagree that no-one cares about me	g2	Likert scale (1 to 5/0 to 4)	No missing value
Disagree that I cannot influence developments	g3	Likert scale (1 to 5/0 to 4)	No missing value
The country is going in the right direction	g4	Likert scale (1 to 5/0 to 4)	No missing value
Satisfied with marriage/relationship	f1	Likert scale (1 to 5/0 to 4)	No missing value; Option 6 (no partner) recoded to 3: neutral
Satisfaction: Family	f2	Likert scale (1 to 5/0 to 4)	No missing value
Satisfaction: Time to do things you want to do	f3	Likert scale (1 to 5/0 to 4)	No missing value

Variable	Label	Recoded/rescaled as likert scale or binary	Missing status
Satisfaction: Leisure time	f4	Likert scale (1 to 5/0 to4)	No missing value
Children did not skip meal past year	f5	Binary (0 to 1)	No missing value
Agree people in your community can be trusted	c1	Binary (0 to 1)	Option 3 (Don't know) set as missing at random
Satisfaction: Friends	c2	Likert scale (1 to 5/0 to 4)	No missing value
It is important to look after the environment	c3	Likert scale (1 to 5/0 to 4)	No missing value
Participated in activities of any club	c4	Binary (0 to 1)	No missing value
Health status in the past 4 weeks	h1	Likert scale (1 to 4/0 to 3)	No missing value
Health does not prevent daily work	h2	Likert scale (1 to 4/0 to 3)	No missing value
Health does not prevent usual social activities	h3	Likert scale (1 to 4/0 to 3)	No missing value
Did not fail to get/look for healthcare	h4	Binary (0 to 1)	No missing value
Personally covered by medical aid	h5	Binary (0 to 1)	Option 6 (Don't know) set as missing at random
Satisfaction: Dwelling	d1	Likert scale (1 to 5/0 to 4)	No missing value
Satisfaction with area where live	d2	Likert scale (1 to 5/0 to 4)	No missing value
Dwelling structure made of bricks or concrete	d3	Binary (0 to 1)	No missing value
Dwelling is owned	d4	Binary (0 to 1)	No missing value
Not overcrowded	d5	Binary (0 to 1)	No missing value
Press is free to write/say what it likes	e1	Likert scale (1 to 5/0 to 4)	No missing value
Highest level of education completed	e2	Likert scale (1 to 6/0 to 5)	Option 19 (Don't know) set as missing at random
Have telephone or cellphone	e3	Binary (0 to 1)	No missing value
Household assets: Television	e4	Binary (0 to 1)	No missing value
Household assets: Internet connection	e5	Binary (0 to 1)	No missing value
Satisfaction: money available to respondent?	w1	Likert scale (1 to 5/0 to 4)	No missing value
Satisfaction: Standard of living	w3	Likert scale (1 to 5/0 to 4)	No missing value
Satisfaction: Working conditions in your job	w4	Likert scale (1 to 5/0 to 4)	No missing value
Employment status	w5	Binary (0 to 1)	No missing value
Total monthly household income	w6	Likert scale (1 to 5/ 0 to 4)	Option 18 (Respondent refused) set as missing at random
Not in debt	w7	Binary (0 to 1)	No missing value
Feel safe walking in area during day	s1	Likert scale (1 to 5/ 0 to 4)	No missing value
Feel safe walking in area after dark	s2	Likert scale (1 to 5/ 0 to 4)	No missing value
Feel safe at home	s3	Likert scale (1 to 5/ 0 to4)	No missing value
Crime situation improving	s4	Likert scale (1 to 3/ 0 to 2)	No missing value

Variable	Label	Recoded/rescaled as likert scale or binary	Missing status
Not a victim of crime in the past year	s5	Binary (0 to 1)	No missing value
How long does it take you after leaving home to reach your destination	r1**	Likert scale (1 to 6/0 to 5)	No missing value
Wants to hear from municipality	r4**	Binary (0 to 1)	No missing value
Government has improved quality of life	r5**	Binary (0 to 1)	No missing value

*Variable g1, 'satisfied with life as a whole', was set aside as a reference. **Variables r1, r4, and r5 were added for this analysis

Data Availability The datasets that were analysed in this study, along with related documentation, is available from the DataFirst repository: 2017/18 QoL V dataset: https://doi.org/10.25828/8yf7-9261, 2015/16 QoL IV dataset: https://doi.org/10.25828/w490-a496, 2013/14 QoL III dataset: https://doi.org/10.25828/gn3g-vc93

Code Availability SPSS and R scripts were used to analyse datasets considered in this study by following the methodology described in Sect. 3. The scripts are available on request.

Declarations

Conflicts of interest The authors declare that they have no conflict of interest.

References

Afifi, A., May, S., Donatello, R., & Clark, V. A. (2019). *Practical multivariate analysis* (6th ed.). Chapman and Hall/CRC.

Archer, M., Sharp, R., Stones, R., & Woodiwiss, T. (1999). Critical realism and research methodology. *Alethia, 2*(1), 12–16. https://doi.org/10.1558/aleth.v2i1.12

Behrens, J. T. (1997). Principles and procedures of exploratory data analysis. *Psychological Methods, 2*, 131–160.

Booysen, F. (2002). An overview and evaluation of composite indices of development. *Social Indicators Research, 59*(2), 115–151. https://doi.org/10.1023/A:1016275505152

Brown, T. (2015). *Confirmatory factor analysis for applied research* (2nd ed.). The Guilford Press.

Byrne, B. M. (2012). *Structural equation modeling with Mplus*. Routledge.

Chowdhury, S., & Squire, L. (2006). Setting weights for aggregate indices: An application to the commitment to development index and human development index. *Journal of Development Studies, 43*(5), 761–771. https://doi.org/10.1080/00220380600741904

Coltman, T., Devinney, T. M., Midgley, D. F., & Venaik, S. (2008). Formative versus reflective measurement models: Two applications of formative measurement. *Journal of Business Research, 61*, 1250–1262.

Cooper, J. C. B. (1983). Factor analysis: An overview. *The American Statistician, 37*(2), 141–147.

Cummins, R. A., Woerner, J., Tomyn, A., & Gibson-Prosser, A. (2012). Quality of life in Australia. In K. C. Land, A. C. Michalos, & M. J. Sirgy (Eds.), *Handbook of Social Indicators and Quality of Life Research* (pp. 459–472). Springer.

Daily Maverick (2016) Violent protests rock S. Africa's capital ahead of vote. https://www.dailymaverick.co.za/article/2016-06-21-violent-protests-rock-s-africas-capital-ahead-of-vote/

De Kadt, J., Ballard, R., Cheruiyot, K., Culwick, C., Gotz, G., Hamann, C., Khanyile, S., Maree, G., Matjomane, M., Moore, R., Naidoo, Y., Parker, A., & Sondzaba, S. (2019). Insights from GCRO's Quality of Life V survey (2017/18). Johannesburg: Gauteng City-Region Observatory. Retrieved January 20, 2020, from https://gcro.ac.za/documents/232/2019.08.29.Data_Insights_clean_formatted.pdf

De Kadt, J. & Crowe-Pettersson, K. (2020). GCRO's Quality of Life survey: Sampling workshop briefing document. Johannesburg: Gauteng City-Region Observatory. Retrieved April 4, 2021, from https://www.gcro.ac.za/documents/443/2020.11.19_QoL_sampling_workshop_briefing_document_formatted.pdf

Diamantopoulos, A. (2005). The C-OAR-SE procedure for scale development in marketing: A comment. *International Journal of Research in Marketing, 22*(1), 1–9. https://doi.org/10.1016/j.ijresmar.2003.08.002

Diamantopoulos, A., & Siguaw, J. A. (2006). Formative versus reflective indicators in organizational measure development: A comparison and empirical illustration. *British Journal of Management, 17*, 263–282. https://doi.org/10.1111/j.1467-8551.2006.00500.x

Diamantopoulos, A., & Winklhofer, H. M. (2001). Index construction with formative indicators: An alternative to scale development. *International Journal of Marketing Research, 38*(2), 269–277. https://doi.org/10.1509/jmkr.38.2.269.18845

Edwards, J. R. (2011). The fallacy of formative measurement. *Organizational Research Methods, 14*(2), 370–388. https://doi.org/10.1177/1094428110378369

Estes, R. J. (2019). The social progress of nations revisited. *Social Indicators Research, 144*, 539–574. https://doi.org/10.1007/s11205-018-02058-9

Everatt, D. & Orkin, M. (1993). Growing up tough: A national survey of South African youth. Submitted to the second national conference on 'Marginalised Youth' convened by the Joint Enrichment Project, Broederstroom, March 1993.

Everatt, D. (2017). Quality of Life in the Gauteng City-Region. *South Africa. Social Indicators Research, 130*(1), 71–86. https://doi.org/10.1007/s11205-015-1127-y

Finn, A., & Kayande, U. (2005). How fine is C-OAR-SE? A generalizability theory perspective on Rossiter's procedure. *International Journal of Research in Marketing, 22*(1), 11–21. https://doi.org/10.1016/j.ijresmar.2004.03.001

Gasper, D. (2010). Understanding the diversity of conceptions of well-being and quality of life. *The Journal of Socio-Economics, 39*(3), 351–360. https://doi.org/10.1016/j.socec.2009.11.006

GCRO (Gauteng City-Region Observatory). (2014). Quality of Life Survey III, 2013/14. Version 1. Johannesburg and Cape Town: GCRO & DataFirst. https://doi.org/10.25828/gn3g-vc93

GCRO (Gauteng City-Region Observatory). (2016). Quality of Life Survey IV, 2015/16. Version 1. Johannesburg and Cape Town: GCRO & DataFirst. https://doi.org/10.25828/w490-a496

GCRO (Gauteng City-Region Observatory). (2019). Quality of Life Survey V, 2017/18. Version 1. Johannesburg and Cape Town: GCRO & DataFirst. https://doi.org/10.25828/8yf7-9261

Glaser, B., & Strauss, A. (1967). *Discovery of grounded theory: Strategies for qualitative research*. CA: Sociology Press.

Glatzer, W. (2012). Cross-national comparisons of quality of life in developed nations, including the impact of globalisation. In K. C. Land, A. C. Michalos, & M. J. Sirgy (Eds.), *Handbook of social indicators and quality of life research*. Springer.

Gorski, P. S. (2013). What is critical realism? And why should you care? *Contemporary Sociology: A Journal of Reviews, 42*(5), 658–670. https://doi.org/10.1177/0094306113499533

Greco, S., Ishizaka, A., Tasiou, M., & Torrisi, G. (2019). On the methodological framework of composite indices: A review of the issues of weighting, aggregation, and robustness. *Social Indicators Research, 141*, 61–94. https://doi.org/10.1007/s11205-017-1832-9

Greyling, T. (2013). A composite index of quality of life for the Gauteng city-region: A principal component analysis approach. GCRO Occasional Paper. Johannesburg: Gauteng City-Region Observatory. Retrieved June 20, 2019, from http://www.gcro.ac.za/sites/default/files/reports/occasional_paper_7.pdf

Greyling, T., & Treganna, F. (2017). Construction and analysis of a composite quality of life index for a region of South Africa. *Social Indicators Research, 131*(3), 887–930. https://doi.org/10.1007/s11205-016-1294-5

Greyling, T., & Tregenna, F. (2020). Quality of life: Validation of an instrument and analysis of relationships between domains. *Development Southern Africa, 37*(1), 19–39. https://doi.org/10.1080/0376835X.2019.1684243

Hagerty, M. R., & Land, K. C. (2012). Issues in composite index construction: The measurement of overall quality of life. In K. C. Land, A. C. Michalos, & M. J. Sirgy (Eds.), *Handbook of social indicators and quality of life research* (pp. 181–200). Springer.

Hair, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2021). *A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM)* (3rd ed.). Sage.

Helmes, E., Goffin, R. D., & Chrisjohn, R. D. (1998). Confirmatory factor analysis of the life satisfaction index. *Social Indicators Research, 45*, 371–390. https://doi.org/10.1023/A:1006966419159

Hosken, G. (2018). Municipal services, billing, and waste are Gauteng residents' biggest gripes, Sunday Times, 13 November 2018. Retrieved March 6, 2020 from https://www.timeslive.co.za/news/south-africa/2018-11-13-municipal-services-billing-and-waste-are-gauteng-residents-biggest-gripes

Howard, M. C. (2016). A review of exploratory factor analysis decisions and overview of current practices: What we are doing and how can we improve? *International Journal of Human-Computer Interaction, 32*(1), 51–62. https://doi.org/10.1080/10447318.2015.1087664

Ihsan, N., & Aziz, B. (2019). A multidimensional analysis of quality of life. *Social Indicators Research, 142*(1), 201–227. https://doi.org/10.1007/s11205-018-1903-6

Jarvis, C. B., MacKenzie, S. B., & Podsakoff, P. M. (2003). A critical review of construct indicators and measurement model misspecification in marketing and consumer research. *Journal of Consumer Research, 30*(2), 199–218. https://doi.org/10.1086/376806

Jennings, R. (2012). Survey of surveys: A survey of citizen-based surveys conducted by provincial and local government in Gauteng. GCRO Occasional Paper. Johannesburg: Gauteng City-Region Observatory. Retrieved May 2, 2019, from https://cdn.gcro.ac.za/media/documents/GCRO_Occasional_Paper_3_Jennings_Survey_of_surveys_October_2012.pdf

Kell, D. B., & Oliver, S. G. (2004). Here is the evidence, now what is the hypothesis? *The Complementary Roles of Inductive and Hypothesis-Driven Science in the Post-Genomic Era, BioEssays, 26*, 99–105.

Kenny, A. (2016). Multiple latent variable models: Confirmatory factor analysis. Retrieved 23 March, 2020, from http://davidakenny.net/cm/mfactor.htm

Lakatos, I. (1980). *The methodology of scientific research programmes*. Cambridge University Press.

Land, K. C., Michalos, A. C., & Sirgy, J. S. (2012). Prologue: The development and evolution of research on social indicators and quality of life. In K. C. Land, A. C. Michalos, & M. J. Sirgy (Eds.), *Handbook of social indicators and quality of life research* (pp. 459–472). Springer.

Liu, Y., Millsap, R. E., West, S. G., Tein, J., Tanaka, R., & Grimm, K. J. (2017). Testing measurement invariance in longitudinal data with ordered-categorical measures. *Psychological Methods, 22*(3), 486–506. https://doi.org/10.1037/met0000075.supp

Lourenço, R., Pereira, V., & Costa, H. G. (2019). Analysis of the Better Life Index through a cluster algorithm. *Social Indicators Research, 142*(2), 477–506. https://doi.org/10.1007/s11205-018-1902-7

Lun, G., Holzer, D., Tappeiner, G., & Tappeiner, U. (2006). The stability of rankings derived from composite indicators: Analysis of the 'Il Sole 24 Ore' quality of life report. *Social Indicators Research, 77*(2), 307–331. https://doi.org/10.1007/s11205-005-4505-z

Maggino, F. (2014). Conceptual design of indicators. In: Michalos A.C. [Ed.]. Encyclopedia of Quality of Life and Well-Being Research. Springer. https://doi.org/10.1007/978-94-007-0753-5_513

Maggino, F., & Zumbo, B. D. (2012). Measuring the quality of life and the construction of social indicators. In K. C. Land, A. C. Michalos, & M. J. Sirgy (Eds.), *Handbook of Social Indicators and Quality of Life Research* (pp. 201–238). Springer.

Mazziotta, M., & Pareto, A. (2017). Synthesis of indicators: The composite indicators approach. In F. Maggino (Ed.), *Complexity in Society: From Indicators Construction to their Synthesis* (pp. 159–191). Springer.

Michalos, A., Smale, B., Labonté, R., Muharjarine, N., Scott, K., Guhn, M., Gadermann, A.M., Zumbo, B.D., Morgan, A., Moore, K., Swystun, L., Holden, B., Bernardin, H., Dunning, B., Graham, P., Bernardin, H., Dunning, B., Graham, P., Brooker, A. & Hyman, I. (2011). The Canadian Index of Wellbeing, Technical paper: Canadian index of wellbeing 1.0. https://uwaterloo.ca/canad

ian-index-wellbeing/sites/ca.canadian-index-wellbeing/files/uploads/files/Canadian_Index_of_Wellbeing-TechnicalPaper-FINAL.pdf

Mikulić, J., Kožić, I. & Krešić, D. (2015). Weighting indicators of tourism sustainability: A critical note. *Ecological Indicators, 48*, 312–314. https://doi.org/10.1016/j.ecolind.2014.08.026

Møller, V., & Roberts, B. (2014). South Africa, Quality of Life. In A. C. Michalos (Ed.), *Encyclopedia of Quality of Life and Well-Being Research.* Dordrecht, Netherlands: Springer.

Mungai, K., & Bayat, A. (2019). An overview of trends in depressive symptoms in South Africa. *South African Journal of Psychology, 49*(4), 518–535. https://doi.org/10.1177/0081246318823580

Nardo, M., Saisana, M., Saltelli, A., Tarantola, S., Hoffman, A. & Giovannini, E. (2005). Handbook on constructing composite indicators: Methodology and user guide. OECD Statistics Working Papers, No. 2005/03, OECD Publishing, Paris. https://doi.org/10.1787/533411815016. Accessed 21 March 2020.

National Planning Commission (2012). National Development Plan 2030. Retrieved August 14, 2021, from https://www.gov.za/sites/default/files/gcis_document/201409/ndp-2030-our-future-make-it-workr.pdf

Nicoletti, G., Scarpetta, S. & Boylaud, O. (2000), Summary indicators of product market regulation with an extension to employment protection legislation. OECD Economics Department Working Papers, No. 226, OECD Publishing, Paris. https://doi.org/10.1787/215182844604

OECD. (2008). Handbook on constructing composite indicators: Methodology and user guide. Paris: OECD Publishing. Retrieved April 4, 2019, from https://www.oecd.org/sdd/42495745.pdf

OECD. (2020). OECD Better Life Index. Organisation for Economic Co-operation and Development (OECD). Retreived October 16, 2020, from http://www.oecdbetterlifeindex.org/#/11111111111

Oliver, C. (2011). Critical realist grounded theory. *British Journal of Social Work.* https://doi.org/10.1093/bjsw/bcr064

Orkin, M. (2020). Technical review of the GCRO QoL surveys: Synthesis report. Johannesburg: Gauteng City-Region Observatory. Retrieved February 23, 2021 from https://gcro.ac.za/documents/441/2020.11.20_Synthesis_report_formatted.pdf

Petter, S., Straub, D. W., & Rai, A. (2007). Specifying formative constructs in information systems research. *MIS Quarterly, 31*(4), 623–656. https://doi.org/10.2307/25148814

Plucker, J. A. (2003). Exploratory and confirmatory factor analysis in gifted education : Examples with self-concept data. *Journal for the Education of the Gifted, 27*(1), 20–35. https://doi.org/10.1177/016235320302700103

Rule, S. (2007). Religiosity and quality of life in South Africa. *Social Indicators Research, 81*, 417–434. https://doi.org/10.1007/s11205-006-9005-2

Saltelli, A. (2007). Composite indicators between analysis and advocacy. *Social Indicators Research, 81*, 65–77. https://doi.org/10.1007/s11205-006-0024-9

Shin, D. C. (2008). Monitoring democratic politics, a market economy, and citizen well-being: The South Korea barometer. In V. Møller, D. Huschka, & A. C. Michalos (Eds.), *Barometers of quality of life around the globe: How are we doing.* Springer.

Simonetto, A. (2012). Formative and reflective models: State of the art. *Electronic Journal of Applied Statistical Analysis, 5*(3), 452–457.

South African Police Service (2017). Annual Crime Report, 2016/17. Retrieved August 14, 2021 from https://www.saps.gov.za/about/stratframework/annual_report/2016_2017/gpw_crime_stats_2017.pdf

Stanojević, A., & Benčina, J. (2019). The construction of an integrated and transparent index of wellbeing. *Social Indicators Research, 143*(3), 995–1015. https://doi.org/10.1007/s11205-018-2016-y

Statistics South Africa (2019). Mid-year population estimates 2019. Retrieved November 9, 2020 from https://www.statssa.gov.za/publications/P0302/P03022019.pdf

Statistics South Africa (2020). Inequality trends in South Africa. Retrieved August 14, 2021 from http://www.statssa.gov.za/publications/Report-03-10-19/Report-03-10-192017.pdf

Stekhoven, D. J., & Bühlmann, P. (2012). MissForest — non-parametric missing value imputation for mixed-type data. *Bioinformatics, 28*(1), 112–118. https://doi.org/10.1093/bioinformatics/btr597

Stiglitz, J., Sen, A., & Fitoussi, J. P. (2009). The measurement of economic performance and social progress revisited: Reflections and overview. Sciences Po publications 2009–33, Sciences Po. Retrieved September 4, 2019 from https://econpapers.repec.org/paper/fcedoctra/0933.htm

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Pearson Education.

Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education, 2*, 53–55. https://doi.org/10.5116/ijme.4dfb.8dfd

Tukey, J. W. (1969). Analyzing data: Sanctification or detective work? *American Psychologist, 24*, 83–91.

Verdugo, M. A., Córdoba, L., & Gómez, J. (2005). Spanish adaptation and validation of the Family Quality of Life Survey. *Journal of Intellectual Disability Research, 49*(10), 794–798.

von Stumm, S., Deary, I. J., & Hagger-Johnson, G. (2013). Life-course pathways to psychological distress a cohort study. *British Medical Journal Open*. https://doi.org/10.1136/bmjopen-2013-002772

Vyas-Doorgapersad, S., & Bangani, A. (2020). The implementation of gender equality within the South African Public Service (1994–2019). *Africa's Public Service Delivery and Performance Review*. https://doi.org/10.4102/apsdpr.v8i1.353

Authors and Affiliations

Samy Katumba[1] · **Julia de Kadt[1]** · **Mark Orkin[2]** · **Paul Fatti[3]**

Julia de Kadt
julia.dekadt@gcro.ac.za

Mark Orkin
mark.orkin@gmail.com

Paul Fatti
paulfatti@gmail.com

[1] Gauteng City-Region Observatory, A partnership between the University of Johannesburg, The University of the Witwatersrand, Gauteng Provincial Government and Organised Local Government in Gauteng, Johannesburg, South Africa

[2] Wits/MRC Development Pathways to Health Research Unit, School of Clinical Medicine, University of the Witwatersrand, Johannesburg, South Africa

[3] School of Statistics and Actuarial Science, University of the Witwatersrand, Johannesburg, Gauteng, South Africa