

A Multidimensional Resilience Capacity Index: Combining Composite Scoring With Complex Systems Methods

Stefan Heinz

stefan.heinz@waikato.ac.nz

University of Waikato

Anthony O'Brien

University of Waikato

Michael O'Sullivan

University of Auckland

Cameron Walker

University of Auckland

Paul Rouse

University of Auckland

Matthew Parsons

Massey University

Jesse Whitehead

University of Waikato

Sione Vaka

University of Waikato

John Hiakita

University of Waikato

Research Article

Keywords: resilience, capacity, composite, Self-Organising Map, configurational analysis

Posted Date: September 29th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10674703/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

A composite index of a multidimensional phenomenon quantifies multiple dimensions into one index value; however, it misses the complexity of human experience such as resilience.

We propose a two-lens hybrid resilience-capacity measure. The first lens measures resilience capacity via the Resilience Capacity Index (RCI), a reflective/formative composite construct using the Adjusted Mazziotta-Pareto Index (AMPI). The validated WHO-5 Well-Being Index (WHO-5) is utilised to test the convergent validity of the RCI. Lens 2 uses the machine-learning technique of case-based Self-Organising Map (SOM) to produce typologies across thirteen sub-blocks spanning three domains. A crosswalk integrates the two lenses enabling a comprehensive picture of resilience capacity.

This hybrid approach addresses four research gaps: i) effective weights with an explicit rationale; ii) a hierarchical Sobol sensitivity analysis that uses weighting blocks of variables; iii) a measurement-model-aware validity framework; iv) a method for calculating configurations based on quantitative data inputs without first assigning weights or specifying the combination.

1 Existing approaches and the four-gaps motivation

Individual and community resilience is influenced by a wide range of social processes operating at different levels (Blinded 2025). Our systematic review (Blinded 2026c) identified four common gaps in practice. Firstly, roughly half of the frameworks used unweighted combinations of indicators and gave no clear rationale (or any rationale) as to why particular weights were employed. Secondly, only 19.4% of studies had analysed the sensitivity of their weighting scales. Thirdly, 31% used reliability tests such as internal consistency for formative composites despite COSMIN (COnsensus-based Standards for the selection of health Measurement INstruments) indicating that such reliability measures should not be used with non-reflective constructs (Mokkink et al. 2018). Finally, although the multidimensional resilience literature strongly emphasises conceptualisation via configuration (typologies of vulnerable and resilient households) (Laurien et al. 2020), these are almost universally both inconsistently operationalised and linked to variable-based composites.

This paper presents the theoretical structure and strategy of two complementary lenses to address all four gaps. Lens 1 applies the Adjusted Mazziotta-Pareto Index (AMPI) with a coefficient-of-variation imbalance penalty (Mazziotta and Pareto 2016) to render arbitrary, unweighted scores into effective weights with substantial variance (Gap 1); it also applies Sobol's variance-based sensitivity indices to conduct global sensitivity analysis across hierarchical architecture levels (addressing Gap 2); and adopts dual COSMIN and OECD/JRC methodologies to assess validity levels appropriate to each of the three domains making up the hierarchical pipeline (Gap 3). The configurational lens (Lens 2) applies a Self-Organising Map (SOM) typology derived from the same data matrix, interpreting the results via Schimpf and Castellani's (2020) COMPLEX-IT case-based approach (Gap 4).

We believe that the greatest contribution of this paper lies not only in the development of the two-lens structure itself, which provides variable-based questions, *How much resilience capacity?*, alongside configurational questions, *What kind of resilience capacity?*, but also in the systematic and comparative relationship between them. In a subsequent paper, we utilise this method to determine New Zealand's General Social Survey (GSS) 2021 and 2023 survey results across all administrative levels linked by Statistics NZ's Integrated Data Infrastructure (IDI); a government run data warehouse, where all government level data is linked by an anonymized identifier.

2 Construct and scope

A sequenced design process led to our resilience capacity framework. First, a scoping review highlighted resilience as a mediating pathway between mental health and wellbeing, identifying individual, social and systemic level factors (Blinded, 2025). Second, using a co-design phase, we asked people, aged 18 + in Aotearoa New Zealand, what individual, social, and systemic factors contribute to their resilience (Blinded, 2026a, 2026b). We then explored the IDI for suitable datasets aligned with the multidimensional resilience factors that emerged during our co-design phase. The scoping review identified that the factors could not be mapped to longitudinal data and could only be mapped onto the GSS. Therefore, the dynamic temporal nature of resilience (bounce back over time) could not be tested. We accordingly scoped the construct to resilience capacity. Measuring capacity at a single point in time is possible with cross-sectional data (Kamble et al. 2025; Food and Agriculture Organization of the United Nations 2016). Second, a systematic review of multidimensional resilience assessment frameworks identified four methodological gaps in the usual construction of these composites (Blinded 2026c) which are addressed here. To probe the capacity-to-outcome link under adversity, we test the resilience capacity against diagnosis and mental-health prescription data.

3 Architecture and validation strategy

3.1 Architecture overview

The two lenses draw from the identical standardised input matrix but deliver different analytic approaches to it. We have sketched this layout in Fig. 1. Both lenses are applied to the same indicator matrix: a range of thirty items labelled as reflective, formative or single-item, thus reflecting their theoretic construct. The items are further divided into thirteen sub-blocks spread across three broad domains (Individual, Social and Systemic). Items are selected and placed into sub-blocks over the co-designed phase. The findings have been reported in standalone papers (Blinded 2026a, 2026b).

3.2 Implementation procedure

The following ten steps summarise the procedure from raw input data to the final outputs of both lenses and the crosswalk (Fig. 1):

- Construct the input matrix of 30 items broken down into 13 sub-blocks representing three domains (Individual, Social, Systemic).
- Classify each sub-block as reflective, formative or single item based on Table 1.

- Aggregate the items into a score for each sub-block, represented by the equal-weighted arithmetic mean of the standardised items in each sub-block (Nardo et al. 2008; Mazziotta and Pareto 2013).
- Rescale each sub-block score to the 70–130 range, where each reference-wave goalpost (the value) and deviation (SD) can be defined based on the 2023 wave (Mazziotta and Pareto 2016).
- Apply AMPI within domains across sub-blocks.
- Aggregate the three domain scores into the RCI by simple arithmetic mean following Mazziotta and Pareto (2016, 2022).
- Train a hexagonal-lattice SOM on the input sub-block scores (Kohonen 1982, 2001).
- Calculate the training quality via the average quantisation and topographic error.
- Apply k-means on the SOM codebook vectors to identify configurational clusters (Vesanto and Alhoniemi 2000); select k via silhouette and elbow inspection.
- Cross-tabulate the RCI deciles against the SOM cluster assignments to construct the crosswalk contingency table and its agreement statistics (Cramér’s V, adjusted Rand index).

3.3 Validity strategy

Table 1 summarises, for each sub-block, the construct measured, the source of validity, and the framework criterion under which it is assessed.

Table 1 Validity provenance				
Sub-block	Construct	Validity Source	Evidence	Framework
Whānau (family and extended family) wellbeing	Relational connection (reflective)	Theoretically justified	CFA, α , ω ; co-design	COSMIN: structural + internal consistency
Institutional trust	Trust in public institutions (reflective)	Theoretically justified	CFA, α , ω ; codesign	COSMIN: structural + internal consistency
Housing conditions	Dwelling adequacy (formative)	Theoretically justified	VIF, directionality; co-design	OECD/JRC Steps 4 and 5
Healthcare access	Geographic access (formative)	Theoretically justified	Whitehead et al. thresholds; co-design	OECD/JRC Step 2
Environmental capacity	Green/blue space access (formative)	Theoretically justified	Healthy Location Index (Marek et al. 2021); co-design	OECD/JRC Step 2
Emergency preparedness	Household preparedness (formative)	Theoretically justified	VIF, directionality; co-design	OECD/JRC Steps 4 and 5
Material wellbeing	Material standard of living	Theoretically justified	MWI-9 (Perry 2009); co-design	OECD/JRC Step 2
Self-efficacy, identity expression, belonging, interpersonal trust, emergency housing	Single-item capacity proxies	Theoretically justified	Co-design	OECD/JRC Step 2; not testable at sub-block level
RCI (total)	Resilience capacity (formative hierarchical)	To be tested in subsequent paper	RCI/WHO-5 convergence, Mental health medication dispensing and long-term-condition associations	OECD/JRC Steps 7 and 8
Test-retest, responsiveness	-	Not testable	Repeated cross-sectional design — distinct respondents per wave	COSMIN - declared limitation
Note: CFA = confirmatory factor analysis; α = Cronbach’s alpha; ω = McDonald’s omega; VIF = variance inflation factor.				

3.4 Lens 1: Variable-based

The variable-based lens yields two parallel index scores. The Resilience Capacity Index (RCI) is formulated as a hierarchical formative composite measuring built resilience capacity. It aggregates the three domains (Individual, Social, Systemic), which in turn are first aggregated through a two-stage process at the sub-block level according to the Municipal Administration Quality Index (MAQI) concept of using a two-staged hierarchical structure including the AMPI. Sub-block scores within each domain are merged using the AMPI, which penalises high between-sub-block variation for each unit. These then form the domain scores which are averaged by simple arithmetic.

3.5 Lens 2: Configurational

Following the approach of Schimpf and Castellani (2020), the configurational lens uses the SOM to model the indicator matrix as a hexagonal lattice and applies the k-means method to the cell values (codebook vectors) to create representative types of resilience configurations (e.g. resilient with formal diagnosis, resilient without formal diagnosis etc.). Cases are interpreted based on the COMPLEX-IT case-based methodology (Schimpf and Castellani 2020) that regards SOM cells as codebook prototypes embedded in configurational space, rather than as instances of a discrete cluster type. The configurations are subsequently characterised by the SOM codebook loadings of each cluster centroid across the thirteen sub-blocks, grouped for interpretive convenience under the three conceptual domains (Individual, Social, Systemic).

3.6 Crosswalk

To illustrate the relationship between the variable-based and configurational approaches, the two sets of scores are overlaid as a contingency table cross-tabulating the decile ranks of the RCI by individual SOM cell locations. The latter is enhanced with statistics describing their degrees of overlap/conflict. Where the measures match, such as when units below the tenth RCI percentile consistently fall into SOM cells characterised by the weakest aggregated capacities, both lenses provide mutually coherent readings. Where they diverge, such as when units in a middle-range RCI decile exhibit marked variation across distinct configurations with different characteristics of their capacity, useful insights emerge about how a single resilience score can obscure meaningful differences that are apparent when examining configurations more closely.

4 Lens 1: Variable-based architecture

4.1 Sub-block measurement-model classification

The thirteen sub-blocks are assigned *reflective*, *formative* or *single-item* status before any form of psychometric test is conducted. The measurement-model stipulates both the rule used to aggregate measures within the sub-block and the validation metric used. Reflective sub-blocks group together multiple items driven by the same underlying latent variable: if the latter changes, all items should increase in the same way (Bollen 1989; Edwards and Bagozzi 2000). Two reflective sub-blocks exist within RCI (Table 1). Both comprise six items measuring relational connection to one's wider family (whānau (family) wellbeing) and one's trust in key public institutions (institutional trust), respectively. Low institutional trust is interpreted here as an indicator of institutional performance toward a population rather than an individual deficit (Reid, Cormack and Paine 2019; Krieger 2014).

Formative sub-blocks consist of multiple items that build and comprise the conceptual factor, rather than mirror it (Diamantopoulos and Winklhofer 2001; Bollen and Diamantopoulos 2017). Items indexing a multidimensional phenomenon are conceptual, non-interchangeable elements of that concept. The measures of each item need not be intercorrelated (Bollen and Diamantopoulos 2017). Four such sub-blocks make up part of the RCI (Table 1): i) housing condition sub-block including four items on warmth, dryness, mould and overall living space; ii) emergency preparedness sub-block consists of three items on water supply, food supply and emergency plans; iii) healthcare access sub-block includes two items: general practice (GP) access and pharmacy access; and iv) environmental capacity sub-block consists of two objective indicators of green space and blue space availability.

The sub-blocks healthcare access and environmental capacity warrant special attention. Access to health care uses objective geographical data calculated on precise commercial road network: (a) the time taken to reach the nearest three GPs (inverted: where a shorter distance means better access), measured at the mesh-block level (the smallest geographic unit used by Stats NZ, approx. 30–60 dwellings) and (b) the time taken to reach the nearest pharmacy (inverted: shorter is better access). The rationale behind including the nearest three general practices is due to GP enrolment data being mostly unavailable in the IDI. We choose the nearest three GPs to acknowledge potential over-flow scenarios. The scoring is based on recent Aotearoa New Zealand studies that discuss variation in travel time to services and thresholds informed by those data and approaches (Whitehead et al. 2019, 2020).

Environmental capacity is mapped following the Healthy Location Index approach (Marek et al. 2021), which measures how accessible green space (parks, vegetation, etc.) and blue space (coastlines, rivers and lakes) is at the meshblock level. The green- and blue-space accessibility scores together enter the RCI matrix as a two-indicator formative construct, validated using the dual framework of the OECD/JRC at Step 4 (Supplementary Figure S1).

Single-item sub-blocks pass an indicator through from an existing, established instrument in unchanged format. There are five single-item sub-blocks: self-efficacy, identity expression, sense of belonging, interpersonal trust and emergency housing. Whereas the five single items that will comprise this first part of the index above differ, the Material Wellbeing sub-block comprises nine items from the Material Wellbeing Index (MWI-9) of Stats NZ and Ministry of Social Development (MSD) (Perry 2009), with the index coming into the RCI already composite (hence we inherited its validation). Single-item sub-blocks rely on co-design support, explicit theoretical anchoring, and inherited validity from source-instrument criterion evidence.

4.2 Within-sub-block aggregation

Both reflective and formative sub-blocks aggregate the included items using an equal-weighted arithmetic mean of standardised item measures. Unit-weighted composites are virtually interchangeable with factor-score-weighted composites in terms of accuracy and perform better out-of-sample (Cohen 1990; Bobko et al. 2007). The defining features separating these two sub-types is not the method through which they form a sub-block, but instead, which statistical assumptions and assumptions in specific theories require verification by the researcher. Reflective sub-blocks need testing on assumptions about internal reliability (e.g. alpha, omega; McNeish and Wolf 2020), whereas formative ones do not require proof of scale assumption-sharing but should pass several tests to account for assumptions such as collinearity via Variance Inflation Factor (VIF) for every item within the block, plus the condition of non-interchangeability described above.

4.3 Normalisation and goalposts (OECD Step 5)

To make cross-block comparisons meaningful on the scale (70–130), each sub-block's resulting score is rescaled according to the goalpost methodology used by Mazziotto and Pareto (2016, Eq. 8) and by Cerqua et al. (2025) in MAQI. For sub-block j and score μ and reference-wave standard deviation σ_j ,

$$\min_j = \mu_j - 2\sigma_j, \max_j = \mu_j + 2\sigma_j$$

The rescaled score $s_{i,j}$ is

$$s_{i,j} = \frac{60(s_{i,j} - \min_j)}{(\max_j - \min_j)} + 70$$

These goalposts were calculated once using a reference wave, then applied uniformly across all observed waves to ensure scores across time are on the same meaningful scale. We used a single wave [GSS 2023] for calculating the RCI scores detailed in the subsequent paper to ensure temporal comparability, use of the same wave can be customised, but it would require recalculating all goalposts again for the RCI scores detailed below.

4.4 Stage-1 aggregation: AMPI within domain (OECD Step 6)

Within each of the three domain types (Individual, Social and Systemic), the constituent sub-block scores are summed using AMPI- with a coefficient-of-variation (cv) imbalance penalty, as applied in Mazziotta and Pareto (2016). For a unit $\vec{v}_i$ and rescaled sub-block measures within domain $\vec{v}_j$, let $M_{\vec{v}_i}$ be the mean sub-block score, $SD_{\vec{v}_i}$ the within-unit standard deviation and $cv = \frac{SD}{M}$.

$$\vec{A} \vec{M} \vec{P} \vec{I}_{\vec{v}_i} = \vec{M}_{\vec{v}_i} - \vec{S} \vec{D}_{\vec{v}_i} \cdot \vec{c} \vec{v}_{\vec{v}_i} = \vec{M}_{\vec{v}_i}$$

As the formula suggests, the imbalance penalty applies when a unit scores variably within a domain, reducing their composite score proportionally. It affects units to differing degrees depending on the relative imbalance and also reflects the concepts of holistic health concept such as the Meihana Model (Pitama et al. 2017). This rule implements a non-compensatory aggregation scheme, preventing units from high performance on certain sub-blocks within the domain from counteracting poor performance in others to improve their final score if sub-block measurements vary in any way (even if only variance-squared) relative to the unit average. AMPI penalty occurs once per unit across three different domains: Individual (two sub-blocks), Social (three sub-blocks), Systemic (seven sub-blocks).

4.5 Stage-2 aggregation: simple arithmetic mean across domains

In the second level of aggregation, the three resulting domain composites (Individual, Social and Systemic) are combined via equal weighting through an arithmetic mean (Greco et al. 2019). We adopted this methodology after reference to MAQI (Cerqua et al. 2025) and following Mazziotta and Pareto (2016, 2022), who argue against double-penalising variance once the within-domain cv penalty has been applied (M. Mazziotta, personal communication, May 12, 2026). Therefore, our architecture applies non-compensatory weighting only at the within-domain level (where its use is theoretically warranted because it highlights unique deficiencies), but applies compensatory weighting between domains (where the components are considered to be theoretically interchangeable elements of total resilience).

4.6 Effective weights

Under the equal-weighted aggregation at both aggregation levels, following Mazziotta and Pareto (2016, 2022), each sub-block's contribution to the final RCI is a consequence of how many sub-blocks share its domain (Greco et al. 2019). Every individual sub-block is estimated to contribute about 1/6 of the weight to the final score. Each social sub-block is weighted at around 1/9 of the way towards the total and Systemic sub-blocks are weighted around 1/21. These differing weights up to roughly three-and-a-half times heavier for individual sub-blocks, reflect a design choice that differences in scope necessitate these varying degrees of influence than being an issue to be smoothed over with further aggregation procedures. We believe a weight count-adjusted method would negate the trade-off between broad and specific measures. That variation is highlighted here but will also be assessed empirically using Sobol sensitivity analysis.

5 Lens 2: Configurational architecture (SOM)

The configurational lens operates on the same matrix which feeds Lens 1. We used the sub-block scores as inputs, rather than the raw indicators, so that both lenses operate on the same signal. Cases with any missing sub-block score are excluded from SOM training on a complete-case basis. A hexagonal-lattice SOM (Kohonen 1982, 2001) is trained on the input sub-block matrix such that each unit represented between 10–20 cases in the original sample (Vesanto and Alhoniemi 2000). SOM training quality is validated through calculation of the standard quantisation and topographic error measures (Kohonen 1995; Kiviluoto 1996).

Configurational types are identified through applying k-means to the SOM codebook weights. The number of clusters, k, is identified via both silhouette analysis (Rousseeuw 1987) and elbow-criterion inspection (Thorndike 1953) on the total within-cluster sum-of-squares across a reasonable number of cluster counts, with k selected as the minimum value that yields both meaningful and theoretically defensible clusters (i.e., clusters that correspond to conceptually interpretable configuration types based on the relative loadings of the codebook weights).

We interpret these cluster types following the configurational framework proposed by Schimpf and Castellani (2020), in which cases are understood not as examples of social types or phenomena in themselves, but as points embedded in configurational space. Each identified cluster is characterised both by the average profile configuration across its component codes in the thirteen sub-blocks and the joint distribution of these members within key demographic sectors relevant to sociological inquiry. The framework also explicitly aims to recognise intra-cluster variance rather than treating a cluster as a type which is assumed homogeneous: to the extent that cases grouped together differ systematically on relevant variables or configuration components, this variation within the cluster is viewed as a critical part of the overall configurational picture, rather than just noise or random error. Cluster validity is assessed using bootstrap resampling. First, the SOM training process followed by k-means cluster analysis is repeated 500 times on independent bootstrap samples of the original data. Subsequently, for each of the 500 bootstrap samples, we calculate the adjusted Rand index (ARI) (Hubert and Arabie 1985) of its clustering solution about the clustering solution found on the full sample data.

6 Crosswalk between Lens 1 and Lens 2

6.1 Rationale

The two lenses addressed different method-based questions to the same body of data. Lens 1 calculated a RCI index (a scalar value, suitable for ordinal-scale comparison or threshold interpretation), asking “How much resilience capacity does this individual have?”. Lens 2 categorised units into different types of configurations (a profile to interpret). These methods are supplementary: a unit may have a high RCI, but a configuration that is low in social-domain capacity, just as a unit might receive a middle score, but belong to a configuration type that is either fairly balanced or uneven (high Individual-domain, low Systemic-domain).

6.2 Operationalisation

We mapped these two different views using a $10 \times k$ cross-classification table, RCI deciles against SOM cluster assignments, reporting joint contribution. We calculated two measures of agreement between the two measurements: Cramér’s V and the adjusted Rand index. The Cramér’s V quantifies the association strength of the RCI level and cluster type. Due to low expected counts in cells toward the bottom-left and top-right corners of the 10-decile $\times$ k-cluster cross-tabulations, we compute p-values for Cramér’s V via Monte Carlo simulation (10,000 replicates of the chi-square statistic under the hypothesis of independence) and we report the conventional Cramér’s V (which tends to underestimate it at small sample sizes) alongside the bias-corrected estimate recommended by Bergsma (2013). The adjusted Rand index, provides a measure of overlap if we choose to consider the RCI deciles themselves as separate categorical clustering of units. We additionally apply a residual-decomposition analysis (Friendly 1994) to identify which combinations of RCI level and cluster type depart most strongly from independence.

6.3 Reading disagreement

Methodological convergence, for example, the bottom decile of the RCI falling into the cluster with the lowest collective-domain capacity profile, would reinforce both lenses’ apparent validity, but would not add much to the content interpretation. By contrast, there exist multiple opportunities for reading disagreement. For instance, a unit with a middle-tier RCI score (decile 5) could fall into a cluster with strong Individual-domain capacity, but weak Systemic-domain capacity. Another unit in the same middle-tier RCI could fall into a cluster with the opposite profile. This suggests these units possess identical values when asked the first question, but different positions in configurations for the second question. Using the 10-decile k-cluster table, we map these two different scalar values (first question) with different configurations (second question) to highlight disagreements.

7 Dual-framework validation strategy

Two frames of reference were utilised to achieve the validation: at the block level appropriate to the measurement model (COSMIN for the reflective sub-blocks and OECD/JRC composite indicator standards for the formative blocks plus the overall RCI framework), and at the configurations layer (SOM, see SOM topography, sub-block leave-one-out perturbation, adjusted Rand Index for bootstrap cluster stability, and cross-wave configurational stability) (Supplementary Figure S1). A silhouette index was calculated, which serves as an index rather than a signal for the validity of the model in terms of robustness, for example, if Silhouette width under multiple tested k values is consistently under 0.10, this is empirical evidence for robustness as a continuous trait (versus an archetype dichotomy). A multi-domain validation protocol was applied covering construct, statistical, sensitivity, cross-cultural, criterion, and interpretability validity (Mokkink et al. 2010, 2018; Prinsen et al. 2018; Nardo et al. 2008). An indicator-decomposition variant (Saisana, Saltelli and Tarantola 2005) was considered but not adopted, because it allocates variance to individual indicators rather than to the hierarchical weight blocks that are the structural object of interest here.

Note

We were unable to assess two measurement properties relevant to the COSMIN standards (test-retest reliability and responsiveness), because we have data from distinct respondent sets for each GSS wave. For Sobol sensitivity testing within-domain values are described qualitatively because the cv^2 penalty can push even relatively well-estimated weights to between 0.8 and 1.0 for total-effect values because its methodology of making assumptions about non-compensatory aggregation makes the cv even with a solid estimate for weights appear high in some cases and vice versa (Homma and Saltelli 1996; Sobol 2001).

8 Illustration on NZ-GSS 2021/2023

Using data from the 2021 and 2023 waves of the New Zealand General Social Survey ($n = 3,483$ (2021); $n = 7,866$ (2023)) linked via the IDI and substantively reported elsewhere, we offer an illustration focusing on applying the architecture and validation strategy in full and outputting the diagnostic statistics required by the dual-framework validation strategy. The following example 1 and 2 are based on hypothetical participant values, while example 3 is based on the results of the trained SOM applied to the 2023 GSS wave

8.1 Example 1: a sub-block score calculation

We present here one example: the whānau wellbeing sub-block, a reflective construct (measured by six items on a five-point Likert scale-). Imagining unit i responded 4, 5, 4, 5, 3, 4 on the six items, then its within-sub-block equal-weighted arithmetic mean is 4.17. The mean score on the sub-block in the 2023 reference wave was $\mu = 3.80$ and the SD was $\sigma = 0.65$. Rescaled to the 70–130 range, this unit’s sub-block score is $60 \cdot \frac{4.17 - 2.50}{5.10 - 2.50} + 70 = 108.5$. The reporting-scale equivalent on the 0–100 scale is given by $y' = (y - 70) \cdot \frac{100}{60}$, mapping the [70, 130] interval linearly onto [0, 100]; here, 108.5 maps to 64.2. On the whānau wellbeing sub-block, the unit scores approximately 0.6 standard deviations above the reference-wave mean and this interpretation is directly conveyed by the reporting-scale figure (Fig. 2).

8.2 Example 2: the AMPI cv^2 imbalance penalty

As an example, in the present Social-domain aggregate for the same unit, three sub-blocks are used (whānau wellbeing, sense of belonging and interpersonal trust), combined via the within-domain $AMPI^-$ formula. Suppose two units, B and C, achieve identical arithmetic means on the Social-domain composite, but their within-unit scores distributions differ. Unit B scores 75, 75, 75 on the three sub-blocks. This means the average is 75, standard deviation is 0, $cv = 0$ and thus its $AMPI^- = 75 \cdot (1 - 0^2) = 75.0$. By contrast, Unit C scores 90, 75, 60 on the same sub-blocks. Although its arithmetic average is still 75, the SD is 12.25 and the cv is 0.163. The CV-squared is thus 0.027. Its $AMPI^-$ is therefore $75 \times (1 - 0.027) = 73.0$. So these two units, while identical in mean value, differ in their Social-domain composite by 2.0 points due to the penalty for variation (Fig. 3). This highlights the non-compensatory character of the aggregation. The system views identically average resilience capacity across a domain as having more substance than resilience capacity that has been focused on one sub-block, at the cost of resilience capacity elsewhere. This is because this approach aligns with the holistic-capacity theory of what constitutes social inclusion.

8.3 Example 3: Configurational profile and crosswalk cell

The configurational lens classifies units into SOM-defined clusters whose cluster centres reveal different configurations. Using the trained SOM for NZ-GSS 2023, five clusters were found, with average RCI scores spanning 26.8 (cluster 5, $n = 756$, configurationally vulnerable, concentrating its low scores on emergency-preparedness) up to 59.4 (cluster 4, $n = 2,721$, high configurational capacity). Cluster 4 also has relatively equal sub-block loadings in the individual, social, and systemic resilience capacity domains. While average RCI puts cluster 4 above cluster 5, the critical distinguishing factor revealed by this analysis is the character of the sub-blocks comprising that capacity. The crosswalk cell combines RCI deciles with SOM clusters. In cells where scalar and configurational reading agree, middling capacity (decile 5 RCI) matched with a mid-capacity configuration (cluster 1), the unit is flagged as middling, as expected. Likewise, low capacity (decile 1 RCI) matched with a configurationally vulnerable profile (cluster 5) flagged it as severely vulnerable and unbalanced.

9 Methodological contributions and scope

9.1 Methodological contributions

The methodological contributions of this study are threefold. Firstly, a measurement-model-aware allocation of COSMIN and OECD/JRC validation criteria to a single hierarchical composite measure is performed: psychometric criteria are assigned to the reflective indicators and composite-indicator criteria are assigned to the formative indicators, as opposed to imposing a single framework. This directly addresses documented empirical practice in the field, wherein internal-consistency reliabilities are often applied to formative composites (Blinded, 2026c) and provides a principled rationale for which framework to apply within a mixed measurement-model architecture. Secondly, a hierarchical Sobol global sensitivity analysis is carried out at the weight simplex rather than at the indicator level. Hierarchical Sobol indices computed at the indicator level are potentially misleading, as they conflate weight choice (architectural) and item distribution (measurement). They also ignore the non-compensatory aggregation logic at the heart of the architecture. Per-simplex reporting directly assesses the exact weighting choices the architecture imposes, allows distinguishing cross-domain weighting (the policy-relevant decision for us) from within-domain weighting decisions and frames within-domain Stage 1 sensitivity as an architectural property rather than a success criterion. Thirdly, a configurational-by-quantitative crosswalk treats the variable-based and case-based readings of the same data as mutually informative rather than as competing or substitute methods. Lens 2 (configurational) is operationalised as more than simply a robustness check on Lens 1 (quantitative). The degree and nature of overlap/divergence between the RCI position and SOM cluster membership become the core methodological object of interest, illuminating areas where the single scalar score obscures useful distinctions that the configurational reading captures.

9.2 Scope and portability

The proposed architecture is portable across composite-indicator domains which exhibit three specific structural traits: a mixed reflective-formative-single-item measurement model at the sub-block level, a small number of conceptually distinct domains at the top level and the availability of an external validated criterion measure for convergent validity (WHO-5 in the case of resilience capacity). Settings lacking all of these characteristics will require an alternative architecture. Configurations sharing a subset of these traits will benefit from the components of the current architecture independently. The dual framework validation strategy, specifically, is transferable to any composite formed by a combination of reflective and formative indicators.

9.3 Limitations of the architecture

There are three architectural limitations that arise from its design, rather than from specific applications. The cv^2 penalty (as proposed in $AMPI$) assumes that intra-domain variation (being low on one sub-block and high on another) is intrinsically undesirable compared with an even distribution within the domain (e. g., being average across all sub-blocks). This is a defensible assumption where the underlying meaning favours homogeneity such as resilience capacity (compensatory logic at the sub-block level would mask an important source of structural vulnerability) but not where the assumption is reversed. The fact that equal weighting across the top level Stage 2 indicators gives rise to unequal per-sub-block weightings on account of differences in sub-block cardinality is structural and can be revealed by the Sobol indices but cannot be addressed by the architecture itself. The configurational lens (Lens 2) carries some limitations specific to SOM, including dependence on lattice geometry and the particular value of training hyperparameters used in constructing it. Bootstrapping the entire analysis (although done here partially as a check) is an insufficient response, however, it partially remedies the uncertainty associated with the training hyperparameters.

9.4 Relationship to other composite-indicator literatures

The architecture builds on three existing literatures while carving out distinct contributions. From the literature on composite indicators originating from OECD/JRC (Nardo et al. 2008), the framework of validation appropriate to formative aggregation and the use of a global sensitivity analysis (GSA) are inherited. From MAQI, an extension of $AMPI$ that integrates multiple domains (Cerqua et al. 2025), the structure of intra-domain- $AMPI$ / across-domain-mean

(level 2 average) architecture and reference wave goalpost anchoring are drawn. From the case-based and COMPLEX-IT literatures (Schimpf and Castellani 2020), the configurational Lens 2 of the SOM output and the conceptualisation of intra-cluster variability as meaningful data rather than noise have been incorporated. The novel contribution of this paper is the unification of these varied strands into a single composite architecture and validation framework, enabling an explicit crosswalk that combines the variable-based and configurational lenses into a stereoscopic view of resilience capacity.

10 Conclusion

A novel dual-lens hierarchical composite measure for population resilience capacity is developed, composed of a variable-based dimension and a configurational dimension (SOM, COMPLEX-IT, case-based interpretation). A crosswalk connects the two lenses. A dual-framework validation methodology is applied: COSMIN criteria to reflective indicators and OECD/JRC (Organisation for Economic Co-operation and Development / Joint Research Centre) criteria to formative indicators. Twenty-five tests across six validation domains, plus a hierarchical Sobol global sensitivity analysis at the architecturally natural level, are conducted. The composite indicator is shown to be transferable to settings that mirror its key architectural traits and it addresses four out of four practical gaps in composite-indicator design revealed in the companion systematic review. Key findings of applying the composite indicator to the New Zealand General Social Survey in 2021 and 2023 are reported in the subsequent paper.

Abbreviations

AMPI
Adjusted Mazziotta–Pareto Index
ARI
Adjusted Rand Index
BMU
Best Matching Unit (BMU₁ first, BMU₂ second—best)
CFA
Confirmatory Factor Analysis
COMPLEX
IT—Case—based methodology of Schimpf and Castellani (2020)
COSMIN
COnsensus—based Standards for the selection of health Measurement INstruments
CV
Coefficient of Variation
GP
General Practice GSA—Global Sensitivity Analysis
GSS
General Social Survey (New Zealand)
IDI
Integrated Data Infrastructure (Statistics New Zealand)
JRC
Joint Research Centre (European Commission)
LINZ
Land Information New Zealand (Toitū Te Whenua)
MAQI
Municipal Administration Quality Index
MSD
Ministry of Social Development (New Zealand)
MWI
9—Material Wellbeing Index (nine—item)
OECD
Organisation for Economic Co—operation and Development
PCA
Principal Component Analysis
QE
Quantisation Error
RCI
Resilience Capacity Index
SOM
Self—Organising Map
TE
Topographic Error
VIF
Variance Inflation Factor

Declarations

Author Contribution

S.H. — conceptualisation, methodology, formal analysis, data curation, software, visualisation, writing (original draft), writing (review and editing). A.O. — supervision, methodology, writing (review and editing). M.O. — supervision, methodology, writing (review and editing). C.W. — supervision, methodology, writing (review and editing). P.R. — supervision, methodology, writing (review and editing). M.P. — supervision, writing (review and editing). J.W. — methodology, data curation, writing (review and editing). S.V. — supervision, methodology. J.H. — supervision, methodology. All authors reviewed and approved the final manuscript and agree to be accountable for the work.

Acknowledgement

The authors would like to thank Prof. Brian Castellani for his feedback on the manuscript; and Dr. Matteo Mazziotta for his input and insight on the Adjusted Mazziotta–Pareto Index. The researchers also thank the Te Tira Maurikura researchers, the early-career research group of Te Pūnaha Matatini, the Aotearoa New Zealand Centre of Research Excellence (CoRE) for complex systems.

Data Availability

The data are held in Statistics New Zealand’s Integrated Data Infrastructure (IDI) and cannot be deposited publicly; access is governed by Statistics New Zealand (Stats NZ). Access to the IDI is restricted to approved researchers. Further information on accessing the study data and code can be obtained by contacting the authors.

References

Note: The authors’ own works are cited as “Blinded” and their reference details have been withheld for double-anonymous review. These citations will be restored upon acceptance.

1. Bergsma, W.: A bias-correction for Cramér’s V and Tschuprow’s T. J. Korean Stat. Soc. 42(3), 323–328 (2013). <https://doi.org/10.1016/j.jkss.2012.10.002>
2. Blinded: Title and source removed for double-anonymous review (2025)
3. Blinded: Title and source removed for double-anonymous review (2026a)
4. Blinded: Title removed for double-anonymous review. Manuscript submitted for publication (2026b)
5. Blinded: Title removed for double-anonymous review. Manuscript under review (2026c)
6. Bobko, P., Roth, P.L., Buster, M.A.: The usefulness of unit weights in creating composite scores: a literature review, application to content validity, and meta-analysis. Organ. Res. Methods 10(4), 689–709 (2007). <https://doi.org/10.1177/1094428106294734>
7. Bollen, K.A.: Structural Equations with Latent Variables. Wiley, New York (1989)
8. Bollen, K.A., Diamantopoulos, A.: In defense of causal-formative indicators: a minority report. Psychol. Methods 22(3), 581–596 (2017). <https://doi.org/10.1037/met0000056>
9. Cerqua, A., Giannantoni, C., Zampollo, F., Mazziotta, M.: The Municipal Administration Quality Index: the Italian case. Soc. Indic. Res. 177, 345–378 (2025). <https://doi.org/10.1007/s11205-024-03511-8>
10. Cohen, J.: Things I have learned (so far). Am. Psychol. 45(12), 1304–1312 (1990)
11. Diamantopoulos, A., Winklhofer, H.M.: Index construction with formative indicators: an alternative to scale development. J. Mark. Res. 38(2), 269–277 (2001). <https://doi.org/10.1509/jmkr.38.2.269.18845>
12. Edwards, J.R., Bagozzi, R.P.: On the nature and direction of relationships between constructs and measures. Psychol. Methods 5(2), 155–174 (2000). <https://doi.org/10.1037/1082-989X.5.2.155>
13. Food and Agriculture Organization of the United Nations: RIMA-II: Resilience Index Measurement and Analysis II. FAO, Rome (2016). <https://www.fao.org/3/a-i5665e.pdf>
14. Friendly, M.: Mosaic displays for multi-way contingency tables. J. Am. Stat. Assoc. 89, 190–200 (1994). <https://doi.org/10.2307/2291215>
15. Greco, S., Ishizaka, A., Tasiou, M., Torrisi, G.: On the methodological framework of composite indices: a review of the issues of weighting, aggregation, and robustness. Soc. Indic. Res. 141, 61–94 (2019). <https://doi.org/10.1007/s11205-017-1832-9>
16. Homma, T., Saltelli, A.: Importance measures in global sensitivity analysis of nonlinear models. Reliab. Eng. Syst. Saf. 52(1), 1–17 (1996). [https://doi.org/10.1016/0951-8320\(96\)00002-6](https://doi.org/10.1016/0951-8320(96)00002-6)
17. Hubert, L., Arabie, P.: Comparing partitions. J. Classif. 2, 193–218 (1985). <https://doi.org/10.1007/BF01908075>
18. Kamble, P.A., Mehta, A., Rani, N.: Measuring multidimensional financial resilience: an ex-ante approach. Soc. Indic. Res. 176(2), 533–567 (2025). <https://doi.org/10.1007/s11205-024-03476-8>

19. Kiviluoto, K.: Topology preservation in self-organizing maps. In: Proceedings of International Conference on Neural Networks (ICNN'96), vol. 1, pp. 294–299 (1996). <https://doi.org/10.1109/ICNN.1996.548907>
20. Kohonen, T.: Self-organized formation of topologically correct feature maps. *Biol. Cybern.* 43(1), 59–69 (1982). <https://doi.org/10.1007/BF00337288>
21. Kohonen, T.: *Self-Organizing Maps*. Springer, Berlin (1995)
22. Kohonen, T.: *Self-Organizing Maps*, 3rd edn. Springer, Berlin (2001). <https://doi.org/10.1007/978-3-642-56927-2>
23. Krieger, N.: Discrimination and health inequities. *Int. J. Health Serv.* 44(4), 643–710 (2014). <https://doi.org/10.2190/HS.44.4.b>
24. Laurien, F., Hochrainer-Stigler, S., Keating, A., Campbell, K., Mechler, R., Czajkowski, J.: A typology of community flood resilience. *Reg. Environ. Change* 20(1), 24 (2020). <https://doi.org/10.1007/s10113-020-01593-x>
25. Marek, L., Hobbs, M., Wiki, J., Kingham, S., Campbell, M.: The good, the bad, and the environment: developing an area-based measure of access to health-promoting and health-constraining environments in New Zealand. *Int. J. Health Geogr.* 20(1), 16 (2021). <https://doi.org/10.1186/s12942-021-00269-x>
26. Mazziotta, M., Pareto, A.: Methods for constructing composite indices: one for all or all for one? *Riv. Ital. Econ. Demogr. Stat.* 67(2), 67–80 (2013)
27. Mazziotta, M., Pareto, A.: On a generalized non-compensatory composite index for measuring socio-economic phenomena. *Soc. Indic. Res.* 127(3), 983–1003 (2016). <https://doi.org/10.1007/s11205-015-0998-2>
28. Mazziotta, M., Pareto, A.: Composite indices construction: the performance interval approach. *Soc. Indic. Res.* 161(2–3), 541–551 (2022). <https://www.jstor.org/stable/48697290>
29. McNeish, D., Wolf, M.G.: Thinking twice about sum scores. *Behav. Res. Methods* 52(6), 2287–2305 (2020). <https://doi.org/10.3758/s13428-020-01398-0>
30. Mokkink, L.B., Terwee, C.B., Patrick, D.L., Alonso, J., Stratford, P.W., Knol, D.L., Bouter, L.M., de Vet, H.C.W.: The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Qual. Life Res.* 19, 539–549 (2010). <https://doi.org/10.1007/s11136-010-9606-8>
31. Mokkink, L.B., de Vet, H.C.W., Prinsen, C.A.C., Patrick, D.L., Alonso, J., Bouter, L.M., Terwee, C.B.: COSMIN Risk of Bias checklist for systematic reviews of patient-reported outcome measures. *Qual. Life Res.* 27, 1171–1179 (2018). <https://doi.org/10.1007/s11136-017-1765-4>
32. Saisana, M., Saltelli, A., Tarantola, S.: Uncertainty and sensitivity analysis techniques as tools for the quality assessment of composite indicators. *J. R. Stat. Soc. Ser. A* 168, 307–323 (2005). <https://doi.org/10.1111/j.1467-985X.2005.00350.x>
33. Perry, B.: *Non-income Measures of Material Wellbeing and Hardship: First Results from the 2008 New Zealand Living Standards Survey, with International Comparisons*. Ministry of Social Development, Wellington (2009)
34. Pitama, S.G., Bennett, S.T., Waitoki, W., Haitana, T.N., Valentine, H., Pahina, J., Taylor, J.E., Tassell-Matamua, N., Rowe, L., Beckert, L., Palmer, S.C., Huria, T.M., Lacey, C.J., McLachlan, A.: A proposed hauora Māori clinical guide for psychologists: using the hui process and Meihana model in clinical assessment and formulation. *N. Z. J. Psychol.* 46(3), 7–19 (2017)
35. Prinsen, C.A.C., Mokkink, L.B., Bouter, L.M., Alonso, J., Patrick, D.L., de Vet, H.C.W., Terwee, C.B.: COSMIN guideline for systematic reviews of patient-reported outcome measures. *Qual. Life Res.* 27, 1147–1157 (2018). <https://doi.org/10.1007/s11136-018-1798-x>
36. Reid, P., Cormack, D., Paine, S.J.: Colonial histories, racism and health—the experience of Māori and Indigenous peoples. *Public Health* 172, 119–124 (2019). <https://doi.org/10.1016/j.puhe.2019.03.027>
37. Rousseeuw, P.J.: Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *J. Comput. Appl. Math.* 20, 53–65 (1987). [https://doi.org/10.1016/0377-0427\(87\)90125-7](https://doi.org/10.1016/0377-0427(87)90125-7)
38. Saisana, M., Saltelli, A., & Tarantola, S. (2005). Uncertainty and sensitivity analysis techniques as tools for the quality assessment of composite indicators. *Journal of the Royal Statistical Society A*, 168, 307–323. <https://doi.org/10.1111/j.1467-985X.2005.00350.x>
39. Schimpf, C., Castellani, B.: COMPLEX-IT: a case-based modelling and scenario simulation platform for social inquiry. *J. Open Res. Softw.* 8(1) (2020). <https://doi.org/10.5334/jors.298>
40. Sobol, I.M.: Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Math. Comput. Simul.* 55(1–3), 271–280 (2001). [https://doi.org/10.1016/S0378-4754\(00\)00270-6](https://doi.org/10.1016/S0378-4754(00)00270-6)
41. Thorndike, R.L.: Who belongs in the family? *Psychometrika* 18(4), 267–276 (1953). <https://doi.org/10.1007/BF02289263>
42. Vesanto, J., Alhoniemi, E.: Clustering of the self-organizing map. *IEEE Trans. Neural Netw.* 11(3), 586–600 (2000). <https://doi.org/10.1109/72.846731>
43. Whitehead, J., Pearson, A.L., Lawrenson, R., Atatoa-Carr, P.: Spatial equity and realised access to healthcare—a geospatial analysis of general practitioner enrolments in Waikato, New Zealand. *Rural Remote Health* 19(4), 5349 (2019). <https://doi.org/10.22605/RRH5349>
44. Whitehead, J., Pearson, A.L., Lawrenson, R., Atatoa-Carr, P.: Defining general practitioner and population catchments for spatial equity studies using patient enrolment data in Waikato, New Zealand. *Appl. Geogr.* 115, 102137 (2020). <https://doi.org/10.1016/j.apgeog.2019.102137>

Figures

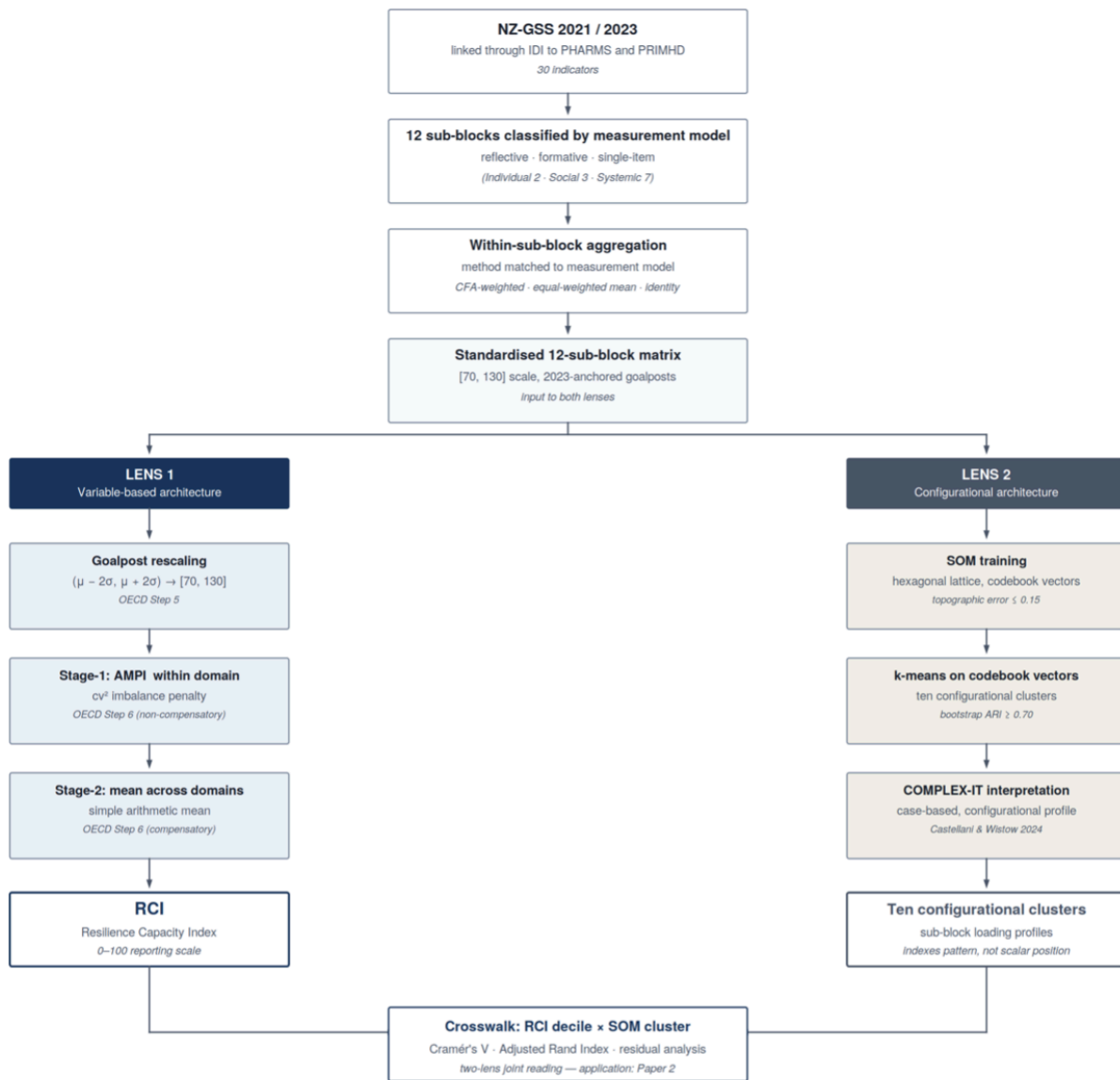


Figure 1

Two lens architecture

Note: IDI = Integrated Data Infrastructure; GSS = General Social Survey; PRIMHD = Programme for the Integration of Mental Health Data; PHARMS = the national Pharmaceutical Collection (pharmaceutical dispensing data).

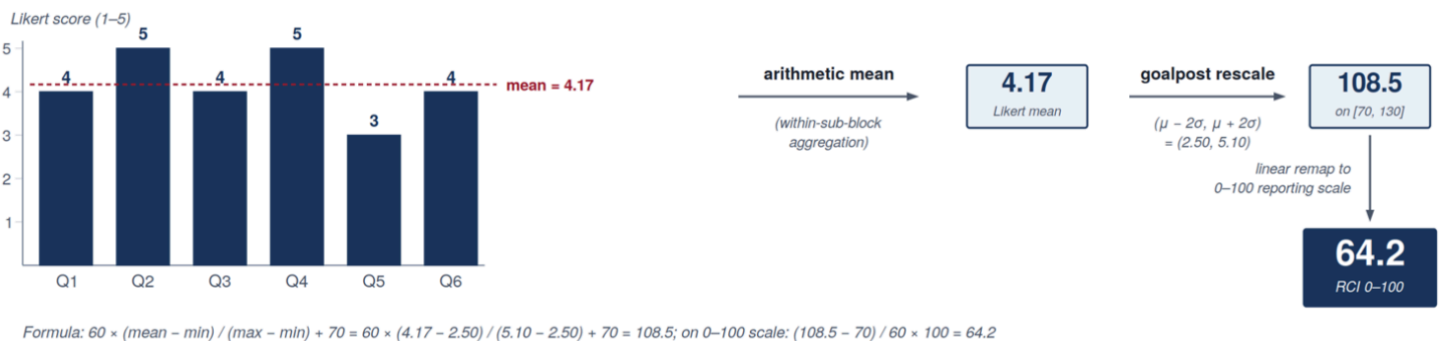


Figure 2

Example: Sub-block calculation

Note: Q1 = Question 1 etc.

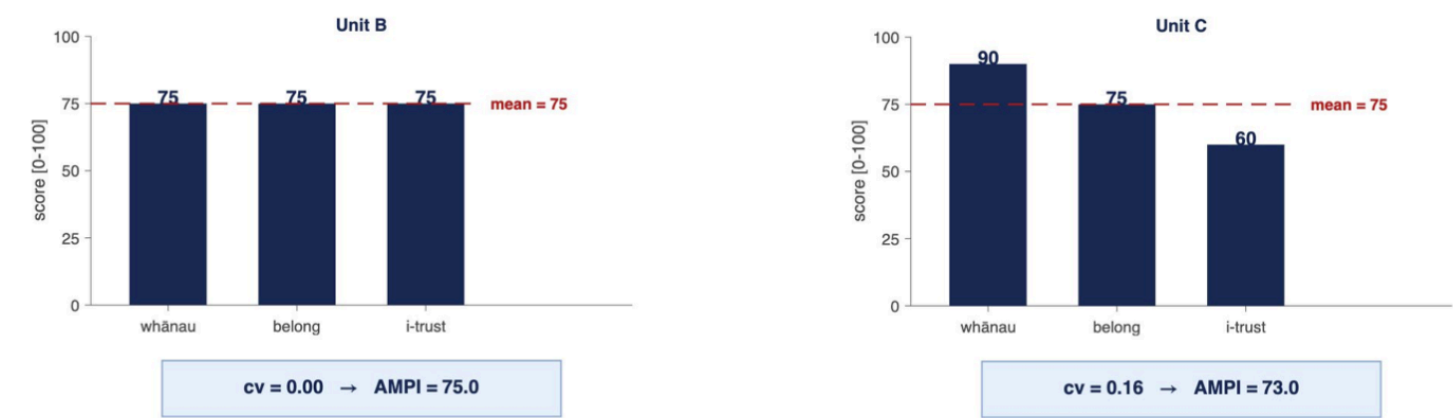


Figure 3

Example: AMPI Penalty

Note: Whānau = Family, extended family; belong = sense of belonging; i-trust = interpersonal trust.

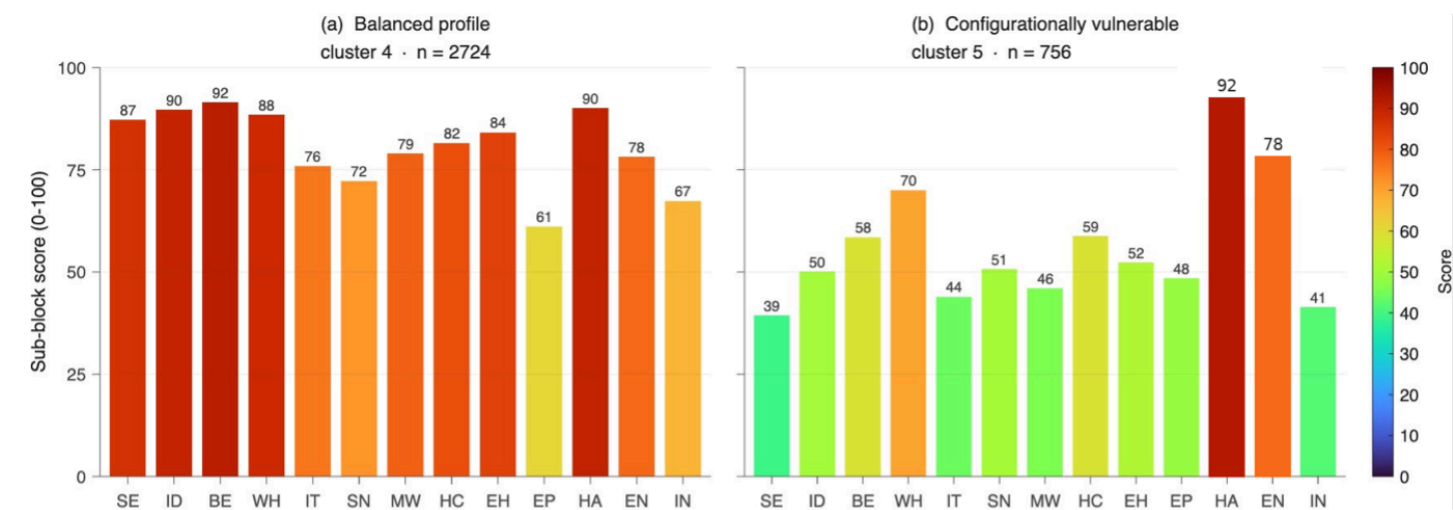


Figure 4

Configurational contrast

Note: SE = self-efficacy; ID = identity expression; BE = belonging; WH = whānau wellbeing; IT = interpersonal trust; SN = social networks; MW = material wellbeing; HC = housing condition; EH = emergency housing; EP = emergency preparedness; HA = healthcare access; EN = environmental capacity; IN = institutional trust.